

## JOURNAL

Submitted (03-DEC-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

## REMARK

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
Contact: MGC help desk

COMMENT

Email: cgapsb-r@mail.nih.gov  
Tissue Procurement: ARCC  
cDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www.shgc.stanford.edu>  
Contact: (Dickson, Mark) [mdc@paxil.stanford.edu](mailto:mdc@paxil.stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MCC clone distribution information can be found through the I.M.A.G.E. Consortium/INTL at: <http://image.llnl.gov>  
Series: TRAK Plate: 20 Row: 1 Column: 3  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 16507953.

## FEATURES

```
source
1. .2621
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:9539 IMAGE:3847282"
/tissue_type="Colon, adenocarcinoma"
/clone_id="NIR MGC_65"
/lab_host="DH10B"
/notes="vector: pCMV-SPORT6"
178. .2400
IDS
```

## CDS

BASE COUNT	563 a	782 c	857 g	419 t
ORIGIN				

Query Match	84.2%	Score 351;	DB 9;	Length 2621;
Best Local Similarity	100.0%	Pred. No. 9.1e-58;		
Matches 351;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 1 TTGGGGGAACCTGTGCGCGAGGGACTCCGCACTCAGGGAACCTGAGACAGAGACAATTGGG 60

Db 178 ATGGGGGAACCTGTGCGCGAGGGACTCCGCACTCAGGGAACCTGAGACAGAGACAATTGGG 237

QY 61 GAGATGATGAGAGAGCCAGCGCCACAGGATGCGTACGTCATCTGCGCCAGCGCCTCAAC 120

Db 238 GAGATGATGAGAGAGCCAGCGCCACAGGATGCGTACGTCATCTGCGCCAGCGCCTCAAC 297

QY 121 CCCCACTCGCCCAAGGCCAAGGTGCTGTGCGACGTCGAGCGAGAGAGAAGTGTGTGCAAGC 180

Db 298 CCCCACTCGCCCAAGGCCAAGGTGCTGTGCGACGTCGAGCGAGAGAGAAGTGTGTGCAAGC 357

QY 181 CCCCGGCTCACCAACAGGCCAATGGGGCCGGGCACTTGTGGATTGTGTGAAGACTGA 240

Db 358 CCCCGCTACCAAGGCGCATGCGGGCCGGGCACTTGCTGATTTGCTGAAAGACTCGA 417

Dy 241 GGGAGAAAGGGGCGCATGCTTCTCTGAGAGCTGAAGTTCCACAAACCTGACGTTC 300

418 GGGAGAAAGGGGCGCATGCTTCTCTGAGAGCTGAAATTCCAAACCTGACGTTC 477

Dy 301 ACCCTGGTACCGGCTGCAGCCTGATGTTCATTCACTTAATTTAGCGGT 351  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 478 ACCCTGGTACCGGCTGCAGCCTGATGTTCATTCACTTAATTTAGCGGT 528

RESULT 2	
AY032927	
LOCUS	3682 bp mRNA linear PRI 07-DEC-2000
DEFINITION	Homo sapiens CARD-containing MAGUK 2 protein (CARMMA2) mRNA,
	complete cds.
ACCESSION	AY032927
VERSION	AY032927.1 GI:14192722
KEYWORDS	.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE  
1 (pages 1 to 3682)  
AUTHORS  
Gaide, O., Martillon, F., Micheau, O., Bonnet, D., Thome, M. and

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

2 (bases 1 to 3682)  
Martinon, F., Thome, M. and Tschopp, J.  
Direct Submission  
Submitted (23-APR-2001) Institute of Biochemistry, University of  
Lausanne, Ch. des Boveresses 155, Epalinges, VD 1066, Switzerland

## FEATURES

```
source
1. .3682
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
gene
1. .3682
  /gene="CARMA2"
CDS
4. .3018
```

BASE COUNT  
ORIGIN

751 a 1100 c 1194 g 637 t

BASE COUNT  
ORIGIN



```

/product="caspase recruitment domain protein 14"
/protein_id="AA53403.1"
/db_xref="GI:12382775"
/translation="MEGLCRDPSALTALDEETLMMESHRRIVRCICPSRLTYLR
OAKVLCOLDEEVLHSPRLTNSAMAGHLLDLTKRGKGAIAFLESIKENHPDYTL
VNGLDOPDPSNPGSGMETSUKTECAAGISLOEELNOKKQKVLARCCOQLEHL
GLAETRAEGLHLEADHSHKREVSSEHFEHLKODMLSLHSLYNSMLQEKELASR
CRSLQSELYLKOELQKRAMVSSCELELOSLKRAVDSESDDELNLKENEKTLRS
LTFSLAEKDLLEQSLDEARGROELVERHSIRERAVALEROREYMEKEQTLQFO
KSMACOLYREKYNALQAOVCELOKEROADYARPSAOSESQSLVSKOSJROYVEL
TPOVCELRQLOLAEPBGVKKOARTEPPCRKORLYRNHAIQPRDNDSCIVSS
TESQLSDLSATSSRELVDSFRSSSPAPPSQSLKRVADDEGSEPMFSSTCLPEP
DPCALPGKAGDPHLDVLELDPADLPQLESLSQVSPGSLDVSESGVLMRRPARIL
SOVTMAFOGDLLEQISVIGNLGFIPIHRYTPGADQMLRPTQIVMVDYASR
PLEKVALDETTLLEAVGLRVDGFCISLKVNTDQYKLLQDLAKVATSGDSFYR
VNLAMBRAGKELOVHCNEVLAVTDMFOGCGCMHAHRNASTMTDIAHGTPYVSR
AQQQLIALIQDTQOCTVTRKSSSGPQCLVAVISMDKAKASPILSLSPRQDLPSRM
EGSSTCFMABSCITLVPTLVWPHRPARPVLVRAVRAKSLSEKLLQGFKCLLA
EYLSQSEYAMSORGDIIOGEVSGGRCWVTHAVALSMERNHALDLVDLSVCTH
RMDIPPIVHVSVMERAKKLKKGLQRLGTSEELRLEAARQEGDLDRAPCLYSLAP
DQMSLDGLSLCVRQALADEQKKVWVTESPR"
273..533
/region="Region: CARD domain"
/misc_feature
588..1433
/region="Region: coiled-coil domain"
/misc_feature
1914..2165
/region="Region: PDZ domain"
/misc_feature
2232..2438
/region="Region: SH3 domain"
/misc_feature
2646..3203
/region="Region: GUK domain"
BASE COUNT 799 a 1187 c 1263 g 682 t
ORIGIN
Query Match 84.2%; Score 351; DB 9; Length 3931;
Best Local Similarity 100.0%; Pred. No. 8.8e-58;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGGGAACCTGTGCGCCAGGGAAGTCCGCACTACGCGCACTGACGAGAGACACTGTGG 60
DB 207 ATGGGGGAACCTGTGCGCCAGGGAAGTCCGCACTACGCGCACTGACGAGAGACACTGTGG 266
QY 61 GAGATGATGAGAGAGCCAGCCGCAAGATGTAAGTGTGCTGCTGCGCCAGCCGCTCACC 120
DB 267 GAGATGATGAGAGAGCCAGCCGCAAGATGTAAGTGTGCTGCTGCGCCAGCCGCTCACC 326
QY 121 CCTTACCTGCGCAGGCGCAAGTGTGCTGCGCAGTGAAGAGAGAGAGAGTGTGCAAGC 180
DB 327 CCTTACCTGCGCAGGCGCAAGTGTGCTGCGCAGTGAAGAGAGAGAGTGTGCAAGC 386
QY 181 CCCCAGCTCAGCAACAGCGCCATGCGGCGCGGCACTTGTGATTTGTGAAGACTCGA 240
DB 387 CCCCAGCTCAGCAACAGCGCCATGCGGCGCGGCACTTGTGATTTGTGAAGACTCGA 446
QY 241 GGGAGAGAGCGGGCGCATGCGCTTCTGAGAGAGCTGAAGTTCCACAACCTGACGTCTAC 300
DB 447 GGGAGAGAGCGGGCGCATGCGCTTCTGAGAGAGCTGAAGTTCCACAACCTGACGTCTAC 506
QY 301 ACCCTGCTCACCGGGCTGACGCGGATGATGACTTCACTTAAGCGGT 351
DB 507 ACCCTGCTCACCGGGCTGACGCGGATGATGACTTCACTTAAGCGGT 557

```

```

REFERENCE 1 (bases 1 to 3772)
AUTHORS McAllister-Lucas,L.M., Inohara,N., Lucas,P.C., Rutland,J.,
Benito,A., Li,Q., Chen,S., Chen,F.F., Yamaoka,S., Verma,I.M.,
Mak,T.W. and Nunez,G.
TITLE Bim1, a MAPK family member linking protein kinase C activation to
JOURNAL Bcl-2-mediated NF-kappaB induction
MEDLINE J. Biol. Chem. 276 (33), 30589-30597 (2001)
PUBMED 11387339
REFERENCE 2 (bases 1 to 3772)
AUTHORS McAllister-Lucas,L.M., Lucas,P.C., Inohara,N., Chen,S., Chen,F.F.
and Nunez,G.
TITLE Direct Submission
JOURNAL Submitted (22-MAR-2001) Department of Pathology, University of
Michigan Medical School, 1500 E. Medical Center Dr., Ann Arbor, MI
48109, USA
FEATURES
source
Location/Qualifiers
1..3772
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1..3772
/gene="Bim2"
304..3303
/gene="Bim2"
/codon_start=1
/product="Bim2"
/protein_id="AA60137.1"
/db_xref="GI:14326097"
/translation="MAELCRMDSTLTALDEEMLMDMLSHRCRIVOSICPSRLPYLR
OAKVLCOLDEEVLHSPRLTNSAMAGHLLDLTKRGKGAIAFLESIKENHPDYTL
VNGLDOPDPSNPGSGMETSUKTECAAGISLOEELNOKKQKVLARCCOQLEHL
GLAETRAEGLHLEADHSHKREVSSEHFEHLKODMLSLHSLYNSMLQEKELASR
CRSLQSELYLKOELQKRAMVSSCELELOSLKRAVDSESDDELNLKENEKTLRS
LTFSLAEKDLLEQSLDEARGROELVERHSIRERAVALEROREYMEKEQTLQFO
KSMACOLYREKYNALQAOVCELOKEROADYARPSAOSESQSLVSKOSJROYVEL
TPOVCELRQLOLAEPBGVKKOARTEPPCRKORLYRNHAIQPRDNDSCIVSS
TESQLSDLSATSSRELVDSFRSSSPAPPSQSLKRVADDEGSEPMFSSTCLPEP
DPCALPGKAGDPHLDVLELDPADLPQLESLSQVSPGSLDVSESGVLMRRPARIL
SOVTMAFOGDLLEQISVIGNLGFIPIHRYTPGADQMLRPTQIVMVDYASR
PLEKVALDETTLLEAVGLRVDGFCISLKVNTDQYKLLQDLAKVATSGDSFYR
VNLAMBRAGKELOVHCNEVLAVTDMFOGCGCMHAHRNASTMTDIAHGTPYVSR
AQQQLIALIQDTQOCTVTRKSSSGPQCLVAVISMDKAKASPILSLSPRQDLPSRM
EGSSTCFMABSCITLVPTLVWPHRPARPVLVRAVRAKSLSEKLLQGFKCLLA
EYLSQSEYAMSORGDIIOGEVSGGRCWVTHAVALSMERNHALDLVDLSVCTH
RMDIPPIVHVSVMERAKKLKKGLQRLGTSEELRLEAARQEGDLDRAPCLYSLAP
DQMSLDGLSLCVRQALADEQKKVWVTESPR"
BASE COUNT 912 a 1023 c 1125 g 712 t
ORIGIN
Query Match 63.5%; Score 264.6; DB 10; Length 3772;
Best Local Similarity 84.6%; Pred. No. 3.6e-41;
Matches 297; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
QY 1 ATGGGGGAACCTGTGCGCCAGGGAAGTCCGCACTACGCGCACTGACGAGAGACACTGTGG 60
DB 304 ATGGGGGAACCTGTGCGCCAGGGAAGTCCGCACTACGCGCACTGACGAGAGAGACTGTGG 363
QY 61 GAGATGATGAGAGAGCCAGCCGCAAGATGTAAGTGTGCTGCTGCGCCAGCCGCTCACC 120
DB 364 GAGATGATGAGAGAGCCAGCCGCAAGATGTAAGTGTGCTGCTGCGCCAGCCGCTCACC 423
QY 121 CCTTACCTGCGCAGGCGCAAGTGTGCTGCGCAGTGAAGAGAGAGAGTGTGCAAGC 180
DB 424 CCTTACCTGCGCAGGCGCAAGTGTGCTGCGCAGTGAAGAGAGAGTGTGCAAGC 483
QY 181 CCCCAGCTCAGCAACAGCGCCATGCGGCGCGGCACTTGTGATTTGTGAAGACTCGA 240
DB 484 TCCCGTTTACCAACAGTGCATGAGAGTGGGCACTTGTGATTTGTGAAGACTCGA 543
QY 241 GGGAGAGAGCGGGCGCATGCGCTTCTGAGAGAGCTGAAGTTCCACAACCTGACGTCTAC 300

```

Db	544	GGGAAGATGAGGCAATTCCTTTCTGGAAGCTGGAAGTTCACAAACCCGTAGTGTAC	603
Db	301	ACCTGTGTCACCGGCTGACCGCTGATTTGACTTGACTTAACCTTTAGCGGT	351
Db	604	ACCTGTGTCACCTGGAGCTGACAGTTCGACATTCGACCTTCAGCAGCTTCAGCGGT	654
RESULT 6	BC029102		
LOCUS	BC029102		
DEFINITION	Mus musculus caspase recruitment domain family, member 14, mRNA (CDNA clone MGC:28122 IMAGE:3379883), complete cds.		
ACCESSION	BC029102		
VERSION	BC029102.1		
KEYWORDS	GI:22137687		
SOURCE	Mus musculus		
ORGANISM	Mus musculus (house mouse)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 3995) Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins P.S., Wagner L., Sherman C.M., Schuller G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P., Datchenko L., Marusina K., Farmer A.A., Rubin G.J., Hong L., Steeleon M., Soares M.B., Bonaldo M.F., Casavant T.L., Schreier T.E., Brownstein M.J., Ueda T.B., Toshitsuki S., Carimni P., Prange C., Rana S.S., Loquillano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan R.J., Malek J.A., Gunaratne P.H., Richards S., Willey K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Vittalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bonfard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz R., Myers R.M., Butterfield Y.S., Krzyzanski M.I., Skalka U., Smolius D.E., Scherch A., Schein J.E., Jones S.J., and Marra M.A.		
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.		
MEDLINE	22388257		
PUBMED	12477932		
REFERENCE	2 (bases 1 to 3995)		
AUTHORS	Strausberg R.		
TITLE	Direct Submission		
JOURNAL	Submitted (01-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: <a href="http://mgc.ncl.nih.gov">http://mgc.ncl.nih.gov</a>		
COMMENT	Contact: MGC help desk Email: <a href="mailto:cgabs-remail.nih.gov">cgabs-remail.nih.gov</a> Tissue Procurement: Gilbert Smith, Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E Consortium (ILNI) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: <a href="http://www.hgsc.bcm.tmc.edu/cdna/">http://www.hgsc.bcm.tmc.edu/cdna/</a> Contact: <a href="mailto:amg@bcm.tmc.edu">amg@bcm.tmc.edu</a> Gunaratne P.H., Garcia A.M., Lu X., Hulyk S.W., Louisseged H., Kowls C.R., Sneed A.J., Martin R.G., Muzny D.M., Nanavati A.N., Gibbs R.A.		
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E Consortium/ILNI at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Series: IRK Plate: 36 Row: e Column: 1 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein. Location/Qualifiers 1..3995		

/organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="MGC:28122 IMAGE:397863"  
 /tissue type="Mammary tumor. Met1loctionien-19F alpha model. 10 month old virgin mouse. Taken by biopsy."  
 /clone\_1ib="NCI\_CGAP\_Mam1"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"  
 1. 3395  
 /gene="Card14"  
 /note="Synonyms: Bimp2, CARMA2"  
 /db\_xref="LocusID:170720"  
 /db\_xref="MGI:2386258"  
 325..3324  
 /codon\_start=1  
 /product="caspase recruitment domain family, member 14"  
 /protein\_id="AAH29102.1"  
 /db\_xref="GI:22137688"  
 /db\_xref="LocusID:170720"  
 /translation="MAELCNDSTLRPLDEPMMDMLSHRCRIYVSGISRLPLPIR  
 QAVYLQDDEEILHSSRFNTNSARVGHLLDLKARKKNALAFLELKHNDPVYLL  
 VTGSLIDISTSGLMETSRLTECLAAGLSLOEIAQKAEVLLRCCQIKERL  
 GLAEAHAGKROLEFVGRMKREKSTHFEVLKIDEMNLSTHYSALKEKELAIR  
 CHSLQELVYKQELORASIVSGCERSRERSIKMSNLSEPOGELNLRBNKILRS  
 MTSIVKEDILBESLDBARSKEPBLDRHLSLBRVAAARQKQVTEHEKQTLQRR  
 KITVDCLELYEKMTMLQGVAELOKEDQVYADRAQMETISQRLVCKDLRRVPEL  
 TERYELMTQRLRLQAEAPGPGKQAEAGBELCRGQRLVRMAVCPDDSDSLSS  
 TBSRLMDLNTSREOMDFRSPSPMPPOQSLYKRVADFLDPESLTFPEVLEMR  
 LQATVDDPTDLPEMDLGDALQSDTDSIQSSRLSNVSESVPRVRRPARKLISGV  
 TVLAFQDALLBQIVYIGNLTGIFTHRTVPGSAADMLRPQTQIMVDYKPKBSL  
 RATLENTTEQAVGLLRVNGSCYLSKLVITEGKXNLIQDLDAVNVSGSFYRVVLD  
 AMORGDEGLQICHNDLHTYDTPFQSRCSMAHNAVMPYTMKDEPITRYSGAQQQ  
 LLALIOMTQRCVPRKPPGPGQKLVIVSVDRKAAVPLTSSFQSDSGDKSGEGRSS  
 VCMRSSECTPLAPYTLVHPHRPARPRPVLPVPLVNGRIILKSLCILQFGFQSCAEYLS  
 QEYATVTSQKGLDIIGESISIGDHMTIRHAYVESLMMNSTALLDVRLDSTRVLRMDM  
 PFIITHSVNKTAKRKGLRKLGRLEGSSEDFLEVARQEBLDRVPLYSLLADSNMS  
 DLSLSLCTVLAALADEKQKVTWTESPC"  
 BASE COUNT 977 a 1095 c 1160 g 763 t  
 ORIGIN  
 Query Match 63.5%; Score 264.6; DB 10; Length 3995;  
 Best Local Similarity 84.6%; Pred. No. 3.6e-41;  
 Matches 297; Conservative 0; Mismatches 54; Indels 0; Gaps 0;  
 1 ATGGGGGAACCTGTGTCGCCGAGGAGACTCCGCACTCAGCGCACTGACGAGAGACACTGTGG 60  
 325 ATGGCAACAACCTGTGTCGCCGAGGAGACTCCGCACTCAGCGCACTGACGAGAGATCCTCTGG 384  
 61 GAGATGATGAGAGAGCCACCGCCACAGAGATGATGACGTGCACTGGCCCAAGCGGCTCAAC 120  
 385 GATATGCTGAGAGATCCAGATGAGAGATTGTGAGAGACATCTGCCCTAGCGGGCTCAAC 444  
 121 CCTCACTGCGCCAGGSCCAAGTGTCTGTGCAGCTGAGACGAGAGAGAGATCTGTCAAGC 180  
 445 CCTCACTGCGCCAGGSCCAAGTGTCTGTGAGAGATGAGAGAGAGAGATCTGTCAAGC 504  
 181 CCCCGGCTCAACAACAAGCGGCATGCGGGCCGGGCACTTGTCTGATTTGCTGAAGACTGGA 240  
 505 TCCGGTTTCAACAACAATGCGCATGAGAGTTGGGCACTTGTGATTCGTAAAGGCCGGA 564  
 241 GGAAGAAGACGGGSCATATGCGCTTCTGTGAAGAGCTGAAGTTCCAAACCTGACGTCTAC 300  
 565 GGAAGAAGATGAGACATGCTTCTGTGAAGAGCTGAAGTTCCAAACCTGACGTCTAC 624  
 301 ACCCTGATCAACGGGCTGACAGCTGATGTTGACTTCAAGTAACTTAACGGT 351  
 625 ACCCTGATCAACGGGCTGACAGCTGATGATTAAGTCAAGCTTCAACGGG 675



LOCUS AC132816 55173 bp DNA linear HTG 04-SEP-2002  
 DEFINITION Homo sapiens chromosome 17 clone RP13-593A24 map 17, LOW-PASS  
 SEQUENCE SAMPLING.  
 AC132816  
 VERSION AC132816.1 GI:22711605  
 KEYWORDS HTG; HTGS PHASE0.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 55173)  
 Authors Birren, B., Nusbaum, C. and Lander, E.  
 TITLE Homo sapiens chromosome 17, clone RP13-593A24  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 55173)  
 Authors Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,  
 Barta, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,  
 Camarata, J., Chang, J., Chazaro, B., Choquet, Y., Collymore, A.,  
 Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,  
 Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,  
 Gaidyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,  
 Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,  
 Karas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,  
 Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,  
 McCarty, M., Meldrum, J., Menus, L., Mihova, T., Mlenga, V.,  
 Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H.,  
 O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,  
 Plunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,  
 Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P.,  
 Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,  
 Tsafaye, S., Theodore, J., Toppan, K., Travers, M., Vassiliev, H.,  
 Vail, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zaitoun, J.,  
 Zembek, L., Zimmer, A. and Zody, M.  
 DIRECT SUBMISSION  
 Submitted (04-SEP-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 All repeats were identified using RepeatMasker:  
 Smit, A. P. A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence.submissions@genome.wi.mit.edu  
 ----- Project Information  
 Center project name: L28179  
 Center clone name: 593\_A\_24

\* NOTE: This record contains 70 individual  
 \* sequencing reads that have not been assembled into  
 \* contigs. Runs of N are used to separate the reads  
 \* and the order in which they appear is completely  
 \* arbitrary. Low-pass sequence sampling is useful for  
 \* identifying clones that may be gene-rich and allows  
 \* overlap relationships among clones to be deduced.  
 \* However, it should not be assumed that this clone  
 \* will be sequenced to completion. In the event that  
 \* the record is updated, the accession number will  
 \* be preserved.

1  
 693: contig of 692 bp in length  
 793: gap of 100 bp  
 1479: contig of 687 bp in length  
 1480: gap of 100 bp  
 1580: contig of 692 bp in length  
 2272: gap of 100 bp  
 2372: contig of 685 bp in length  
 3057: gap of 100 bp  
 3157: gap of 100 bp  
 3858: contig of 702 bp in length  
 3859: gap of 100 bp  
 3959: gap of 100 bp  
 4632: contig of 674 bp in length  
 4732: gap of 100 bp  
 4733: contig of 686 bp in length  
 5418: gap of 100 bp  
 5419: gap of 100 bp

5519: contig of 681 bp in length  
 6199: gap of 100 bp in length  
 6200: contig of 670 bp in length  
 6300: gap of 100 bp in length  
 6970: contig of 710 bp in length  
 7070: gap of 100 bp in length  
 7780: contig of 689 bp in length  
 8569: gap of 100 bp in length  
 8669: contig of 636 bp in length  
 9365: gap of 100 bp in length  
 9465: contig of 707 bp in length  
 10172: gap of 100 bp in length  
 10272: contig of 692 bp in length  
 10964: gap of 100 bp in length  
 11064: contig of 688 bp in length  
 11751: gap of 100 bp in length  
 11852: contig of 696 bp in length  
 12547: gap of 100 bp in length  
 12647: contig of 662 bp in length  
 12648: gap of 100 bp in length  
 13310: contig of 100 bp in length  
 13410: gap of 100 bp in length  
 14099: contig of 689 bp in length  
 14198: gap of 100 bp in length  
 14699: contig of 671 bp in length  
 14870: gap of 100 bp in length  
 14970: contig of 638 bp in length  
 15668: gap of 100 bp in length  
 15767: contig of 712 bp in length  
 15768: gap of 100 bp in length  
 16479: contig of 689 bp in length  
 16579: gap of 100 bp in length  
 16880: contig of 689 bp in length  
 17369: gap of 100 bp in length  
 18039: contig of 670 bp in length  
 18139: gap of 100 bp in length  
 18330: contig of 691 bp in length  
 18829: gap of 100 bp in length  
 18929: contig of 703 bp in length  
 19332: gap of 100 bp in length  
 19732: contig of 700 bp in length  
 20433: gap of 100 bp in length  
 20533: contig of 681 bp in length  
 21213: gap of 100 bp in length  
 21314: contig of 675 bp in length  
 21988: gap of 100 bp in length  
 21989: contig of 713 bp in length  
 22088: gap of 100 bp in length  
 22801: contig of 674 bp in length  
 22802: gap of 100 bp in length  
 22901: contig of 674 bp in length  
 23575: gap of 100 bp in length  
 23576: contig of 680 bp in length  
 24355: gap of 100 bp in length  
 24455: contig of 691 bp in length  
 25146: gap of 100 bp in length  
 25246: contig of 701 bp in length  
 25247: gap of 100 bp in length  
 25947: contig of 653 bp in length  
 26047: gap of 100 bp in length  
 26740: contig of 692 bp in length  
 26741: gap of 100 bp in length  
 27543: contig of 703 bp in length  
 27643: gap of 100 bp in length  
 27644: contig of 695 bp in length  
 28338: gap of 100 bp in length  
 28339: contig of 699 bp in length  
 28439: gap of 100 bp in length  
 29137: contig of 687 bp in length  
 29237: gap of 100 bp in length  
 29924: contig of 687 bp in length  
 30024: gap of 100 bp in length  
 30712: contig of 688 bp in length  
 30812: gap of 100 bp in length  
 30713: contig of 692 bp in length  
 31504: gap of 100 bp in length  
 31604: contig of 718 bp in length  
 32322: gap of 100 bp in length  
 32422: contig of 682 bp in length  
 33104: gap of 100 bp in length  
 33204: contig of 695 bp in length  
 33899: gap of 100 bp in length  
 33900: contig of 691 bp in length  
 34690: gap of 100 bp in length

*	34791	34790: gap of 100 bp	in length
*	34792	34791: contig of 691 bp	in length
*	35582	35581: gap of 100 bp	in length
*	35582	35581: contig of 696 bp	in length
*	36278	36277: gap of 100 bp	in length
*	36378	36372: contig of 695 bp	in length
*	37073	37122: gap of 100 bp	in length
*	37173	37856: contig of 684 bp	in length
*	37857	37956: gap of 100 bp	in length
*	37857	38652: contig of 696 bp	in length
*	38653	38752: gap of 100 bp	in length
*	38753	39442: contig of 688 bp	in length
*	39441	39540: gap of 100 bp	in length
*	39541	40218: contig of 678 bp	in length
*	40219	40318: gap of 100 bp	in length
*	40319	41015: contig of 697 bp	in length
*	41016	41115: gap of 100 bp	in length
*	41116	41890: contig of 705 bp	in length
*	41821	41920: gap of 100 bp	in length
*	41921	42600: contig of 680 bp	in length
*	42601	42700: gap of 100 bp	in length
*	42701	43366: contig of 686 bp	in length
*	43387	43486: gap of 100 bp	in length
*	43487	44279: contig of 693 bp	in length
*	44180	44279: gap of 100 bp	in length
*	44280	44939: contig of 660 bp	in length
*	44480	45039: gap of 100 bp	in length
*	45040	45703: contig of 664 bp	in length
*	45704	45803: gap of 100 bp	in length
*	45804	46470: contig of 667 bp	in length
*	46471	46570: gap of 100 bp	in length
*	46571	47264: contig of 694 bp	in length
*	47265	47364: gap of 100 bp	in length
*	47365	48069: contig of 705 bp	in length
*	48070	48159: gap of 100 bp	in length
*	48170	48847: contig of 678 bp	in length
*	48848	48947: gap of 100 bp	in length
*	48948	49636: contig of 689 bp	in length
*	49637	49766: gap of 100 bp	in length
*	49737	50437: contig of 701 bp	in length
*	50438	50537: gap of 100 bp	in length
*	50538	51213: contig of 676 bp	in length
*	51214	51313: gap of 100 bp	in length
*	51314	51999: contig of 686 bp	in length
*	52000	52099: gap of 100 bp	in length
*	52100	52792: contig of 693 bp	in length
*	52793	52892: gap of 100 bp	in length
*	52893	53594: contig of 702 bp	in length
*	53595	53694: gap of 100 bp	in length
*	53695	54377: contig of 683 bp	in length
*	54378	54477: gap of 100 bp	in length
*	54478	55173: contig of 696 bp	in length

FEATURES  
source

Query Match	50.8%;	Score 212;	DB 2;	Length 55173;
Best Local Similarity	100.0%;	Pred. No. 3.5e-31;		
Matches 212;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	ATGGGGGAACTGTGCCCCGAGGAACTCCGCACTCAGCACTGGAACAGAGGAACTGTGG	60
Db	51604	ATGGGGGAACTGTGCCCCGAGGAACTCCGCACTCAGCACTGGAACAGAGGAACTGTGG	51665
QY	61	GAGATGATGAGAGAGCCACCGCCACAGGATGCTACGTGCATTCGCCCAAGCCGCTCAC	120
Db	51664	GAGATGATGAGAGAGCCACCGCCACAGGATGCTACGTGCATTCGCCCAAGCCGCTCAC	51722
QY	121	CCCTACCTGCGCCAGGCCCAAGTGTCTGTGCAGCTGGAACAGAGAGAGTGTCTGCACAG	180
Db	51724	CCCTACCTGCGCCAGGCCCAAGTGTCTGTGCAGCTGGAACAGAGAGAGTGTCTGCACAG	51783
QY	181	CCCCGCGCTCACCAACAGCGCCATGCGGAGCCGG	212

Db 51784 CCCCCGCTCACCAACAGCGCCATGCGGCGG 51815

LOCUS	AC123764	14400 bp	DNA	linear	PRI 11-MAR-2003
DEFINITION	Homo sapiens chromosome 17, clone CTC-784N20, complete sequence.				
ACCESSION	AC123764				
VERSION	AC123764.18	GI:28913166			
KEYWORDS	HTG.				
SOURCE	Homo sapiens (human)				

REFERENCE	1 (bases 1 to 144000)
AUTHORS	Birren, B., Nusbaum, C. and Lander, E.
TITLE	Homo sapiens chromosome 17, clone CTC-784N20
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 144000)
AUTHORS	Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,

Chazairo, B., Choepel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A., Cooke, P., Dearliano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., FitzHugh, M., Gage, D., Gabaian, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Jumaeson, R., Jones, C., Kamat, A., Karaas, A., Kells, C., Labrecque, K., Lamares, R., Landers, T., Lehoczyk, U., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menais, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norby, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Punnamang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talama, J., Testaye, S., Theodore, J., Totham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, K., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, D., Zimmer, A., and Zody, M.

**TITLE**  
Submitted (01-JUN-2002) Whitehead Institute/MIT Center for Genome  
Direct Submissions  
**JOURNAL**  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (Dates 1 to 14400)  
**REFERENCE**  
Birtan, B., Nushbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,  
AUTHORS

COOK, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,  
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,  
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafsz, N.,  
Hagoe, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,  
Lindblad-Toh, K., Liu, G., Maclean, C., MacDonald, P., Major, J.,  
Matthews, C., McCarthy, M., Meldrum, J., Meneses, L., Minova, T.,  
Miyaga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., North, C.,  
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,  
Peterson, K., Pumphuang, P., Pierre, L., Raymond, C., Retta, R.,  
Reese, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupbach, R.,  
Saman, S., Severy, P., Smith, J., Roy, A., Spencer, B., Stange-Thomann, N.,  
Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K.,  
Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,  
Wyman, D., Young, G., Zainoun, J., Zemek, L., Zimmer, A. and Zody, M.  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
Submitted (16-JAN-2003) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
4 (bases 1 to 14400)  
Birten, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,  
Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,  
Bogunlavsky, L., Boukhgalter, B., Camarata, L., Chang, Y., Choquet, Y.,  
Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K.,  
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,  
Ferreira, P., Fitzgerald, M., Gage, D., Galagan, D., Gardyna, S.,



Db 13002 CCGTACTGCGCAGGCGCAAGGCTGTGTGCTGACCTGACGAGGAGGAGGTGCTGCACAGC 13061

Qy 181 CCCCCGCTCACCAACAGCGCCATGCGGCGCGG 212

Db 13062 CCCCCGCTCACCAACAGCGCCATGCGGCGCGG 13093

RESULT 9  
AC087741/c  
LOCUS AC087741 182016 bp DNA linear PRI 17-DEC-2002  
DEFINITION Homo sapiens chromosome 17, clone RP11-334C17, complete sequence.  
AC087741  
VERSION AC087741.18 GI:27151453  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 182016)  
Birten,B., Nusbaum,C. and Lander,E.  
Homo sapiens chromosome 17, clone RP11-334C17  
2 (bases 1 to 182016)  
Birten,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S., Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Dolly,M., Fero,S., Ferreira,P., Fitzhugh,N., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Karatas,A., Labocque,K., Lamazares,R., Landers,T., Lenoczky,J., Levine,R., Liu,G., Maclean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Mpheters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Sougnaz,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Travers,M., Trivisan,A., Trigliio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (22-JAN-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 182016)  
Birten,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barta,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cooke,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Nord,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (28-NOV-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
4 (bases 1 to 182016)  
Birten,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barta,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,

Canarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cooke,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Nord,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (17-DEC-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Dec 17, 2002 this sequence version replaced gi:25167154.  
All repeats were identified using RepeatMasker:  
Smit,A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RW/RepeatMasker.html>  
----- Genome Center  
Center: Whitehead Institute / MIT Center for Genome Research  
Center code: WMR  
Web site: <http://www-seg.wi.mit.edu>  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: 112382  
Center clone name: 334\_C17  
-----  
FEATURES  
source Location/Qualifiers  
1..182016  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="17"  
/map="17"  
/clone="RP11-334C17"  
/clone\_1lb="RP11-11 Human Male BAC"  
complement(2597..2904)  
/rpt\_family="AluXs"  
complement(3378..3689)  
/rpt\_family="AluY"  
complement(3699..3988)  
/rpt\_family="AluJc"  
complement(4291..4403)  
/rpt\_family="MIR"  
4866..4948  
/rpt\_family="GA-rich"  
5205..5254  
/rpt\_family="L2"  
complement(5534..5695)  
/rpt\_family="AluSp"  
complement(6506..6609)  
/rpt\_family="MER5B"  
6661..6763  
/rpt\_family="MER5A"  
complement(6879..7187)  
/rpt\_family="AluXs"  
7241..7336  
/rpt\_family="MER81"  
7332..7514  
/rpt\_family="MER20"  
8710..8732  
/rpt\_family="GAAA)n"  
complement(8788..8868)  
/rpt\_family="MER20"  
9120..9606  
/rpt\_family="MER1A"  
9637..9943  
/rpt\_family="AluXs"

repeat\_region complement(10266, .10516)  
 /rpt\_family="AluSx"  
 repeat\_region complement(10582, .10738)  
 /rpt\_family="FPM"  
 repeat\_region complement(10806, .10962)  
 /rpt\_family="AluSg"  
 unsure complement(10958, .10973)  
 /note="single clone coverage"  
 repeat\_region 11049, .11236  
 /rpt\_family="AluUb"  
 repeat\_region 12381, .12509  
 /rpt\_family="Charlie8"  
 repeat\_region complement(12611, .12820)  
 /rpt\_family="AluSg"  
 repeat\_region complement(12949, .13219)  
 /rpt\_family="AluSc"  
 repeat\_region complement(15127, .15161)  
 /rpt\_family="L2"  
 repeat\_region complement(15228, .15283)  
 /rpt\_family="MIR"  
 repeat\_region 15284, .15647  
 /rpt\_family="THE1C"  
 repeat\_region complement(15648, .15685)  
 /rpt\_family="MIR"  
 repeat\_region complement(15782, .15860)  
 /rpt\_family="L2"  
 repeat\_region 17124, .17422  
 /rpt\_family="AluSx"  
 repeat\_region complement(17967, .18101)  
 /rpt\_family="L1ME4A"  
 repeat\_region complement(19532, .19840)  
 /rpt\_family="AluSx"  
 repeat\_region complement(21106, .21402)  
 /rpt\_family="AluJo"  
 repeat\_region 22639, .23145  
 /rpt\_family="L1MC4a"  
 repeat\_region 23146, .23203  
 /rpt\_family="AluSx"  
 repeat\_region 23204, .23323  
 /rpt\_family="L1MC4a"  
 repeat\_region 23324, .23637  
 /rpt\_family="AluSx"  
 repeat\_region 23638, .23851  
 /rpt\_family="L1MC4a"  
 repeat\_region 24094, .24625  
 /rpt\_family="L1MC4a"  
 repeat\_region 24825, .25117  
 /rpt\_family="AluJo"  
 repeat\_region complement(25762, .25901)  
 /rpt\_family="L2"  
 repeat\_region 26264, .26442  
 /rpt\_family="TTGGG)n"  
 repeat\_region complement(28083, .28218)  
 /rpt\_family="MIR"  
 repeat\_region 28255, .28430  
 /rpt\_family="MIR"  
 repeat\_region 28783, .28913  
 /rpt\_family="MIR"  
 repeat\_region 29046, .29084  
 /rpt\_family="MER41D"  
 repeat\_region 29111, .29392  
 /rpt\_family="AT\_rich"  
 repeat\_region 29396, .29433  
 /rpt\_family="AluUb"  
 repeat\_region /rpt\_family="CAAA)n"  
 Query Match 50.8%; Score 212; DB 9; Length 182016;  
 Best Local Similarity 100.0%; Pred. No. 3,1e-31;  
 Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 GAGATGATGAGAGCCACCCGACAGAGTGTAGTGCATCTGCCAGCCGCTCACC 120  
 DB 28614 GAGATGATGAGAGCCACCCGACAGAGTGTAGTGCATCTGCCAGCCGCTCACC 28615  
 QY 121 CCTTACCTGCGCCGACGAGGAGTGTGTGCGACGTGACGAGAGGAGTGTGACAGC 180  
 DB 28614 CCTTACCTGCGCCGACGAGGAGTGTGTGCGACGTGACGAGAGGAGTGTGACAGC 28555  
 QY 181 CCCCAGCTCACACACAGGCGCATGCGGCGG 212  
 DB 28554 CCCCAGCTCACACACAGGCGCATGCGGCGG 28523  
 RESULT 10  
 AC015559  
 LOCUS  
 DEFINITION Homo sapiens chromosome 17 clone RP11-334C17, WORKING DRAFT  
 AC015559  
 SEQUENCE, 11 unordered pieces.  
 AC015559  
 VERSION AC015559.10 GI:13173726  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCES  
 AUTHORS  
 1 (bases 1 to 187865)  
 Abola, A.P., Bruno, D., Conn, L., Dela Rosa, M., Faulkner, D.,  
 Federpiet, N., Glukhov, S., Hansen, N., Herman, Z.S., Hyman, R.,  
 Mac, J., Lam, B., Marathe, R., Miranda, M., Morehouse, A.J., Nguyen, M.,  
 Oefner, P., Palm, C.J., Ramirez, D., Southwick, A.M., Wilhelmy, J.,  
 Yu, S., and Davis, R.W.  
 Unpublished  
 2 (bases 1 to 187865)  
 Bruno, D., Conn, L., Dela Rosa, M., Federpiet, N., Foreman, P.,  
 Glukhov, S., Hansen, N., Hyman, R., Mac, J., Marathe, R.,  
 Morehouse, A.J., Oefner, P., Palm, C.J., Ramirez, D., Wilhelmy, J.,  
 Yu, S., and Davis, R.W.  
 Direct Submission  
 Submitted (17-NOV-1999) DNA Sequencing and Technology Center,  
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
 USA  
 On Mar 1, 2001 this sequence version replaced gi:11968315.  
 COMMENT  
 ----- Genome Center  
 Center: Stanford DNA Sequencing and Technology Development  
 Center  
 Center code: SDSSTDC  
 Web site: http://sequence-www.stanford.edu/group/human/  
 Contact: hum-info@sequence.stanford.edu  
 ----- Project Information  
 Center project name: RP11-334C17  
 ----- Summary Statistics  
 Sequencing Vector: M13mp18; X02513  
 Chemistry: Dye-primer; 11% of reads  
 Chemistry: Dye-terminator Big Dye; 89% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 180520 bases at least Q40  
 Consensus quality: 181812 bases at least Q30  
 Consensus quality: 182303 bases at least Q20  
 Insert size: 186865; agarose-fp  
 Quality coverage: 9.0x in Q20 bases; agarose-fp  
 Quality coverage: 9.2x in Q20 bases; sun-of-contigs.  
 NOTE: This is a "working draft" sequence. It currently  
 consists of 11 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.

\* 1 3078: contig of 3078 bp in length  
 \* 3079 3178: gap of unknown length

```

* 3179 6672: contig of 3494 bp in length
* 6673 6772: gap of unknown length
* 6773 10097: contig of 3325 bp in length
* 10098 10197: gap of unknown length
* 10198 15387: contig of 5190 bp in length
* 15388 15487: gap of unknown length
* 15488 24379: contig of 8891 bp in length
* 24379 24479: gap of unknown length
* 24479 34708: contig of 10222 bp in length
* 34708 34807: gap of unknown length
* 34808 46756: contig of 11949 bp in length
* 46757 46857: gap of unknown length
* 46857 67497: contig of 20641 bp in length
* 67498 67598: gap of unknown length
* 67598 91511: contig of 23914 bp in length
* 91512 91611: gap of unknown length
* 91612 118431: contig of 26820 bp in length
* 118432 118531: gap of unknown length
* 118532 187865: contig of 69334 bp in length.

```

## FEATURES

## SOURCE

```

1. 187865
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/clone="RP11-334C17"
/clone_id="RP11 human BAC library 11"
1. 3078
/note="assembly_name:Contig31"
3179. 6672
/note="assembly_name:Contig32"
6773. 10097
/note="assembly_name:Contig33"
10198. 15387
/note="assembly_name:Contig34"
15488. 24378
/note="assembly_name:Contig35"
24479. 34707
/note="assembly_name:Contig36
clone end:SP6"
34808. 46756
/note="assembly_name:Contig37"
46857. 67497
/note="assembly_name:Contig38"
67598. 91511
/note="assembly_name:Contig39"
91612. 118431
/note="assembly_name:Contig40"
118532. 187865
/note="assembly_name:Contig41"
BASE COUNT 44858 a 47815 c 48011 g 46163 t 1018 others
ORIGIN

```

```

Query Match 50.8%; Score 212; DB 2; Length 187865;
Best Local Similarity 100.0%; Pred.No.3,1e-11;
Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 ATGGGGGAACTGTCGCGAGGACTCCGCACTCAAGGCACTGACGAGAGACACTGTGG 60
DB 73791 ATGGGGGAACTGTCGCGAGGACTCCGCACTCAAGGCACTGACGAGAGACACTGTGG 73850
QY 61 GAGTGTATGAGAGACCCGCGACAGAGATGTAAGCTGCACTCCCGACCCGCTTACC 120
DB 73851 GAGTGTATGAGAGACCCGCGACAGAGATGTAAGCTGCACTCCCGACCCGCTTACC 73910
QY 121 CCCACCTGCGCCGCGCGCAAGAGTCTGTGCGACGTGACGAGAGAGAGTGTGACAGC 180
DB 73911 CCCACCTGCGCCGCGCGCAAGAGTCTGTGCGACGTGACGAGAGAGAGTGTGACAGC 73970
QY 181 CCCCGGCTCACCAACAGCGCCATGCGGGCGCG 212
DB 73971 CCCCGGCTCACCAACAGCGCCATGCGGGCGCG 74002

```

```

RESULT 11
AC109321
LOCUS
DEFINITION
Homo sapiens chromosome 17 clone RP11-1079G10 map 17, 5 unordered
pieces.
AC109321
AC109321.15 GI:26655779
HTG: HTGS_PHASE1; HTGS_FULLTOP; HTGS_CANCELLED.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 207822)
2 (bases 1 to 207822)
REFERENCE
Birren,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Bouckgalter,B.,
Brown,A., Camarata,J., Campoliano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kelle,C., Larocque,K., Lamazares,R.,
Landers,T., Lehocsky,J., Levine,R., Liu,G., Maclean,C.,
Macdonald,P., Major,J., Marguis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrum,J., Meneses,D., Minova,T.,
Mlenga,V., Murphy,T., Naylor,U., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunhahang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnpack,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Straus,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.U., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS

```

```

TITLE
JOURNAL
COMMENT

```

```

Direct Submission
Submitted (27-MAR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Dec 13, 2002 this sequence version replaced gi:26291922.
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RW/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: M18R
Web site: http://www-seq.wi.mit.edu

```



*	1	632: contig of 692 bp in length
*	693	792: gap of 100 bp
*	793	1479: contig of 687 bp in length
*	1480	1579: gap of 100 bp
*	1580	2271: contig of 692 bp in length
*	2272	2371: gap of 100 bp
*	2372	3056: contig of 685 bp in length
*	3057	3156: gap of 100 bp
*	3157	3858: contig of 702 bp in length
*	3858	3958: gap of 100 bp
*	3959	4632: contig of 674 bp in length
*	4633	4732: gap of 100 bp
*	4733	5418: contig of 686 bp in length
*	5419	5518: gap of 100 bp
*	5519	6199: contig of 681 bp in length
*	6200	6299: gap of 100 bp
*	6300	6699: contig of 670 bp in length
*	6300	7059: gap of 100 bp
*	6370	7779: contig of 710 bp in length
*	7070	7879: gap of 100 bp
*	7780	8568: contig of 689 bp in length
*	7880	8669: gap of 100 bp
*	8669	9364: contig of 696 bp in length
*	9365	9464: gap of 100 bp
*	9465	10171: contig of 707 bp in length
*	10172	10271: gap of 100 bp
*	10272	10963: contig of 692 bp in length
*	10364	11063: gap of 100 bp
*	11064	11751: contig of 688 bp in length
*	11752	11851: gap of 100 bp

```

* 11852 12547: contig of 696 bp in length
* 12548 12647: gap of 100 bp
* 12648 13309: contig of 662 bp in length
* 13310 13409: gap of 100 bp
* 13410 14098: contig of 689 bp in length
* 14099 14198: gap of 100 bp
* 14199 14869: contig of 671 bp in length
* 14870 14969: gap of 100 bp
* 14970 15667: contig of 698 bp in length
* 15668 15767: gap of 100 bp
* 15768 16479: contig of 712 bp in length
* 16480 16579: gap of 100 bp
* 16580 17268: contig of 689 bp in length
* 17269 17368: gap of 100 bp
* 17369 18038: contig of 670 bp in length
* 18039 18138: gap of 100 bp
* 18139 18829: contig of 691 bp in length
* 18830 18929: gap of 100 bp
* 18930 19632: contig of 703 bp in length
* 19633 19732: gap of 100 bp
* 19733 20432: contig of 700 bp in length
* 20433 20532: gap of 100 bp
* 20533 21213: contig of 681 bp in length
* 21214 21313: gap of 100 bp
* 21314 21988: contig of 675 bp in length
* 21989 22088: gap of 100 bp
* 22089 22801: contig of 713 bp in length
* 22802 22901: gap of 100 bp
* 22902 23575: contig of 674 bp in length
* 23576 23675: gap of 100 bp
* 23676 24355: contig of 680 bp in length
* 24356 24455: gap of 100 bp
* 24456 25146: contig of 691 bp in length
* 25147 25247: gap of 100 bp
* 25248 25947: contig of 701 bp in length
* 25948 26047: gap of 100 bp
* 26048 26740: contig of 693 bp in length
* 26741 26840: gap of 100 bp
* 26841 27543: contig of 703 bp in length
* 27544 27643: gap of 100 bp
* 27644 28338: contig of 695 bp in length
* 28339 28438: gap of 100 bp
* 28439 29137: contig of 699 bp in length
* 29138 29237: gap of 100 bp
* 29238 29924: contig of 687 bp in length
* 29925 30024: gap of 100 bp
* 30025 30712: contig of 688 bp in length
* 30713 30812: gap of 100 bp
* 30813 31504: contig of 692 bp in length
* 31505 31604: gap of 100 bp
* 31605 32322: contig of 718 bp in length
* 32323 32422: gap of 100 bp
* 32423 33104: contig of 682 bp in length
* 33105 33204: gap of 100 bp
* 33205 33899: contig of 693 bp in length
* 33900 33999: gap of 100 bp
* 33999 34690: contig of 691 bp in length
* 34691 34790: gap of 100 bp
* 34791 35481: contig of 691 bp in length
* 35482 35581: gap of 100 bp
* 35582 36277: contig of 696 bp in length
* 36278 36377: gap of 100 bp
* 36378 37072: contig of 695 bp in length
* 37073 37172: gap of 100 bp
* 37173 37856: contig of 684 bp in length
* 37857 37956: gap of 100 bp
* 37957 38652: contig of 696 bp in length
* 38653 38752: gap of 100 bp
* 38753 39440: contig of 688 bp in length
* 39441 39540: gap of 100 bp
* 39541 40218: contig of 678 bp in length
* 40219 40318: gap of 100 bp
* 40319 41015: contig of 697 bp in length

```

## FEATURES

source 1..55173

Query Match 49.2%; Score 205; DB 2; Length 55173;  
 Best Local Similarity 100.0%; Pred. No. 7.9e-30;  
 Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 211 GGGCACTGCTGAGATTGCTGAGACTGAGGAAAGGGGCGCATGCTTCTGAG 270
    |||
DB 20167 GGGCACTGCTGAGATTGCTGAGACTGAGGAAAGGGGCGCATGCTTCTGAG 20108
    |||
QY 271 ACCCTGAAGTTCACAAACCTGACCTTACACCCCTGTCACCGGAGCTGACGCTGATGT 330
    |||
DB 20107 ACCCTGAAGTTCACAAACCTGACCTTACACCCCTGTCACCGGAGCTGACGCTGATGT 20048
    |||
QY 331 GACTTCAGTAAGTTCAGGAGGAGAGCTCCGACTTGAAGGTTGGACGGCACTTTCAG 390
    |||
DB 20047 GACTTCAGTAAGTTCAGGAGGAGAGCTCCGACTTGAAGGTTGGACGGCACTTTCAG 19988
    |||
QY 391 AACCTCAGGCTCTCTGTATACCCAG 415
    |||
DB 19987 AACCTCAGGCTCTCTGTATACCCAG 19963
    |||

```

RESULT 13  
 AC137736 74301 bp DNA linear HTG 28-NOV-2002  
 LOCUS Homo sapiens chromosome 17 clone RPL3-365D24 map 17, LOW-PASS  
 DEFINITION SEQUENCE SAMPLING.  
 AC137736  
 ACCESSION AC137736.1 GI:25815500  
 VERSION AC137736.1  
 KEYWORDS HTG; HTGS; PHASE0.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 74301)

AUTHORS Birren, B., Nusbaum, C. and Lander, E.  
TITLE Homo sapiens chromosome 17, clone RP13-365D24  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 74301)  
AUTHORS Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,  
Barra, N., Bastien, V., Bloom, T., Boguslavsky, L., Bokhgale, B.,  
Cammarata, J., Chang, J., Chazaro, B., Choepel, Y., Collamore, A.,  
Cook, A., Cooke, P., DeArrelano, K., Dewar, K., Diaz, S., Dodge, S.,  
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,  
Gardner, S., Gird, S., Graham, L., Grand-Pierre, N., Hafez, N.,  
Hagos, B., Horton, L., Hulme, M., Iliev, I., Johnson, R., Jones, C.,  
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,  
Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J.,  
Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T.,  
Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,  
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,  
Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R.,  
Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schnupack, R.,  
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,  
Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K.,  
Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,  
Wymann, D., Young, G., Zaitoun, J., Zembek, L., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (28-NOV-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
Smit, A.P.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
----- Genome Center -----  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIMR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)  
----- Project Information -----  
Center project name: L28728  
Center clone name: 365\_D\_24  
-----  
\* NOTE: This record contains 91 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.  
\*  
\* 1 707: contig of 707 bp in length  
\* 708 807: gap of 100 bp  
\* 808 1540: contig of 733 bp in length  
\* 1541 1640: gap of 100 bp  
\* 1641 2344: contig of 704 bp in length  
\* 2345 2444: gap of 100 bp  
\* 2445 3156: contig of 712 bp in length  
\* 3157 3256: gap of 100 bp  
\* 3257 3978: contig of 722 bp in length  
\* 3979 4078: gap of 100 bp  
\* 4079 4781: contig of 703 bp in length  
\* 4782 4881: gap of 100 bp  
\* 4882 5594: contig of 713 bp in length  
\* 5595 5694: gap of 100 bp  
\* 5695 6405: contig of 711 bp in length  
\* 6406 6505: gap of 100 bp  
\* 6506 7197: contig of 692 bp in length  
\* 7198 7297: gap of 100 bp  
\* 7298 8015: contig of 718 bp in length  
\* 8016 8115: gap of 100 bp  
\* 8116 8818: contig of 703 bp in length  
\* 8819 8918: gap of 100 bp  
\* 8919 9653: contig of 735 bp in length  
\* 9654 9753: gap of 100 bp  
\* 9754 10454: contig of 701 bp in length

10455 10554: gap of 100 bp  
10555 11277: contig of 723 bp in length  
11278 11377: gap of 100 bp  
11378 12081: contig of 704 bp in length  
12082 12181: gap of 100 bp  
12182 12913: contig of 732 bp in length  
12914 13014: gap of 100 bp  
13015 13742: contig of 729 bp in length  
13743 13842: gap of 100 bp  
13843 14560: contig of 718 bp in length  
14561 14660: gap of 100 bp  
14661 15383: contig of 722 bp in length  
15384 15482: gap of 100 bp  
15483 16198: contig of 716 bp in length  
16199 16299: gap of 100 bp  
16299 17006: contig of 708 bp in length  
17007 17106: gap of 100 bp  
17107 17817: contig of 711 bp in length  
17818 17917: gap of 100 bp  
17918 18642: contig of 725 bp in length  
18643 18742: gap of 100 bp  
18743 19458: contig of 716 bp in length  
19459 19558: gap of 100 bp  
19559 20264: contig of 706 bp in length  
20265 20364: gap of 100 bp  
20365 21103: contig of 739 bp in length  
21104 21203: gap of 100 bp  
21204 21910: contig of 707 bp in length  
21911 22010: gap of 100 bp  
22011 22731: contig of 721 bp in length  
22732 22831: gap of 100 bp  
22832 23556: contig of 725 bp in length  
23557 23656: gap of 100 bp  
23657 24373: contig of 717 bp in length  
24374 24473: gap of 100 bp  
24474 25166: contig of 693 bp in length  
25167 25266: gap of 100 bp  
25267 25949: contig of 683 bp in length  
25950 26049: gap of 100 bp  
26050 26763: contig of 714 bp in length  
26764 26863: gap of 100 bp  
26864 27583: contig of 720 bp in length  
27584 28421: gap of 100 bp  
28422 28521: contig of 738 bp in length  
28522 29223: gap of 100 bp  
29224 29323: contig of 702 bp in length  
29324 30031: gap of 100 bp  
30032 30131: contig of 708 bp in length  
30132 30859: gap of 100 bp  
30860 30959: contig of 728 bp in length  
30960 31693: gap of 100 bp  
31694 31793: contig of 734 bp in length  
31794 32516: gap of 100 bp  
32517 32616: contig of 723 bp in length  
32617 33338: gap of 100 bp  
33339 33438: contig of 722 bp in length  
33439 34162: gap of 100 bp  
34163 34262: contig of 724 bp in length  
34263 34983: gap of 100 bp  
34984 35082: contig of 720 bp in length  
35083 35785: gap of 100 bp  
35786 35885: contig of 703 bp in length  
35886 36612: gap of 100 bp  
36613 36712: contig of 727 bp in length  
36713 37460: gap of 100 bp  
37461 37560: contig of 748 bp in length  
37561 38293: gap of 100 bp  
38294 38393: contig of 733 bp in length  
38394 39114: gap of 100 bp  
39115 39214: contig of 721 bp in length  
39215 39943: gap of 100 bp  
39944 40043: contig of 725 bp in length

```

* 40044 40763: contig of 720 bp in length
* 40764 40863: gap of 100 bp
* 40864 41562: contig of 699 bp in length
* 41563 41662: gap of 100 bp
* 41663 42386: contig of 724 bp in length
* 42387 42486: gap of 100 bp
* 42487 43202: contig of 716 bp in length
* 43203 43302: gap of 100 bp
* 43303 44007: contig of 705 bp in length
* 44008 44107: gap of 100 bp
* 44108 44794: contig of 687 bp in length
* 44795 44894: gap of 100 bp
* 44895 45581: contig of 687 bp in length
* 45582 45681: gap of 100 bp
* 45682 46363: contig of 682 bp in length
* 46364 46463: gap of 100 bp
* 46464 47166: contig of 703 bp in length
* 47167 47266: gap of 100 bp
* 47267 48002: contig of 736 bp in length
* 48003 48102: gap of 100 bp
* 48103 48832: contig of 729 bp in length
* 48833 48931: gap of 100 bp
* 48932 49651: contig of 720 bp in length
* 49652 49751: gap of 100 bp
* 49752 50488: contig of 737 bp in length
* 50489 50588: gap of 100 bp
* 50589 51295: contig of 707 bp in length
* 51296 51395: gap of 100 bp
* 51396 52123: contig of 728 bp in length
* 52124 52223: gap of 100 bp
* 52224 52935: contig of 712 bp in length
* 52936 53035: gap of 100 bp
* 53036 53743: contig of 708 bp in length
* 53744 53843: gap of 100 bp
* 53844 54539: contig of 695 bp in length
* 54539 54638: gap of 100 bp
* 54639 55354: contig of 716 bp in length
* 55355 55454: gap of 100 bp
* 55455 56159: contig of 705 bp in length
* 56160 56259: gap of 100 bp
* 56260 56986: contig of 727 bp in length
* 56987 57086: gap of 100 bp
* 57087 57807: contig of 721 bp in length

```

Query Match 45.3% Score 188.8; DB 2; Length 74301;  
 Best Local Similarity 62.3% Pred. No. 9.9e-27;  
 Matches 202; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

```

QY 1 ATGGGAGAACTGTCCGAGGAGCTCCGACTCAGGCACTGAGCAGAGAGACACTGTGG 60
Db 24185 ATGGGAGAACTGTCCGAGGAGCTCCGACTCAGGCACTGAGCAGAGAGACACTGTGG 24244
QY 61 GAGATGATGAGAGACCAACCGCCACAGAGATCTGATCTGATCCCGACCGGCTTACC 120
Db 24245 GAGATGATGAGAGACCAACCGCCACAGAGATCTGATCTGATCCCGACCGGCTTACC 24304
QY 121 CCTTACTGCGCCGAGGAGAGAGTGTGCTGCAAGTGAAGAGAGAGAGTGTGACACG 180
Db 24305 CCTTACTGCGCCGAGGAGAGAGTGTGCTGCAAGTGAAGAGAGAGAGTGTGACACG 24364
QY 181 CCGCGGCTCACAACAGGCGCATGCGGCGGCGGCACTTGCTGATTCGAGAGACTCGA 240
Db 24365 CCGCGGCTCANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 24424
QY 241 GGGAGAGACGGGAGCATGCTCTGAGAGAGCTGAAGTTCCACAACCTGACGTAC 300
Db 24425 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 24484
QY 301 ACCCTGTGTACCGGCGCTGACGCT 324
Db 24485 TATCTCGCGCTGTGTGATTCCT 24508

```

RESULT 14  
 AC132197/c  
 LOCUS  
 DEFINITION  
 Homo sapiens chromosome 11 clone RP11-452112 map 11, LOW-PASS  
 SEQUENCE SAMPLING.  
 AC132197  
 AC132197.1 GI:22549815  
 HTG; HTGS\_PHASE0.  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 65608)  
 Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,  
 Birren,B., Baskien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,  
 Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,  
 Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
 Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,  
 Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,  
 Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,  
 Karakas,A., Kelle,C., Landers,T., Levine,R., Lindblad-Toh,K.,  
 Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,  
 McCarthy,M., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V.,  
 Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,  
 O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,  
 Phunhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
 Roman,J., Roy,A., Schauer,S., Schupbach,R., Seaman,S., Severy,P.,  
 Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,  
 Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,  
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,  
 Zembek,L., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (30-AUG-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 All repeats were identified using RepeatMasker:  
 Smit, A.P.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RW/RepeatMasker.html

TITLE  
 JOURNAL  
 COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WISR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 Project Information  
 Center project name: L28114  
 Center clone name: 452\_I\_12

NOTE: This record contains 81 individual  
 sequencing reads that have not been assembled into  
 contigs. Runs of N are used to separate the reads  
 and the order in which they appear is completely  
 arbitrary. Low-pass sequence sampling is useful for  
 identifying clones that may be gene-rich and allows  
 overlap relationships among clones to be deduced.  
 However, it should not be assumed that this clone  
 will be sequenced to completion. In the event that  
 the record is updated, the accession number will  
 be preserved.  
 1  
 688: contig of 688 bp in length  
 689  
 788: gap of 100 bp  
 789  
 1492: contig of 704 bp in length  
 1493  
 1592: gap of 100 bp  
 1593  
 2304: contig of 712 bp in length  
 2305  
 2404: gap of 100 bp  
 2405  
 3111: contig of 707 bp in length  
 3112  
 3211: gap of 100 bp  
 3212  
 3928: contig of 717 bp in length  
 3929  
 4028: gap of 100 bp  
 4029  
 4758: contig of 730 bp in length  
 4759  
 4858: gap of 100 bp

```
* 4859 5578: contig of 720 bp in length
* 5579 5678: gap of 100 bp
* 5679 6395: contig of 717 bp in length
* 6396 6495: gap of 100 bp
* 6496 7212: contig of 717 bp in length
* 7213 7312: gap of 100 bp
* 7313 8028: contig of 716 bp in length
* 8029 8128: gap of 100 bp
* 8129 8857: contig of 723 bp in length
* 8858 8957: gap of 100 bp
* 8958 9662: contig of 705 bp in length
* 9663 9762: gap of 100 bp
* 9763 10469: contig of 707 bp in length
* 10470 10569: gap of 100 bp
* 10570 11275: contig of 706 bp in length
* 11276 11375: gap of 100 bp
* 11376 12101: contig of 726 bp in length
* 12102 12201: gap of 100 bp
* 12202 12918: contig of 717 bp in length
* 12919 13018: gap of 100 bp
* 13019 13710: contig of 692 bp in length
* 13711 13810: gap of 100 bp
* 13811 14524: contig of 714 bp in length
* 14525 14624: gap of 100 bp
* 14625 15348: contig of 724 bp in length
* 15349 15448: gap of 100 bp
* 15449 16130: contig of 682 bp in length
* 16131 16230: gap of 100 bp
* 16231 16904: contig of 674 bp in length
* 16905 17004: gap of 100 bp
* 17005 17701: contig of 697 bp in length
* 17702 17801: gap of 100 bp
* 17802 18518: contig of 717 bp in length
* 18519 18618: gap of 100 bp
* 18619 19331: contig of 713 bp in length
* 19332 19431: gap of 100 bp
* 19432 20131: contig of 700 bp in length
* 20132 20231: gap of 100 bp
* 20233 20933: contig of 702 bp in length
* 20934 21033: gap of 100 bp
* 21034 21743: contig of 710 bp in length
* 21744 21843: gap of 100 bp
* 21844 22563: contig of 720 bp in length
* 22564 22663: gap of 100 bp
* 22664 23381: contig of 716 bp in length
* 23382 23481: gap of 100 bp
* 23482 24183: contig of 702 bp in length
* 24184 24283: gap of 100 bp
* 24284 24979: contig of 696 bp in length
* 24980 25079: gap of 100 bp
* 25080 25806: contig of 727 bp in length
* 25807 25906: gap of 100 bp
* 25907 26601: contig of 695 bp in length
* 26602 27416: contig of 715 bp in length
* 27417 27516: gap of 100 bp
* 27517 28224: contig of 708 bp in length
* 28225 28324: gap of 100 bp
* 28325 29023: contig of 699 bp in length
* 29024 29123: gap of 100 bp
* 29124 29839: contig of 716 bp in length
* 29840 29939: gap of 100 bp
* 29940 30650: contig of 711 bp in length
* 30651 30750: gap of 100 bp
* 30751 31462: contig of 712 bp in length
* 31463 31562: gap of 100 bp
* 31563 32274: contig of 712 bp in length
* 32275 32374: gap of 100 bp
* 32375 33093: contig of 719 bp in length
* 33094 33193: gap of 100 bp
* 33194 33910: contig of 717 bp in length
* 33911 34010: gap of 100 bp
* 34011 34721: contig of 711 bp in length
```

```
* 34722 34821: gap of 100 bp
* 34822 35545: contig of 724 bp in length
* 35546 35645: gap of 100 bp
* 35646 36356: contig of 711 bp in length
* 36357 36456: gap of 100 bp
* 36457 37175: contig of 719 bp in length
* 37176 37275: gap of 100 bp
* 37276 38005: contig of 730 bp in length
* 38006 38105: gap of 100 bp
* 38106 38822: contig of 717 bp in length
* 38823 39922: gap of 100 bp
* 39923 39642: contig of 720 bp in length
* 39643 39742: gap of 100 bp
* 39743 40463: contig of 721 bp in length
* 40464 40563: gap of 100 bp
* 40564 41293: contig of 730 bp in length
* 41294 41393: gap of 100 bp
* 41394 42102: contig of 709 bp in length
* 42103 42202: gap of 100 bp
* 42203 42887: contig of 685 bp in length
* 42888 42987: gap of 100 bp
* 42988 43705: contig of 718 bp in length
* 43706 43805: gap of 100 bp
* 43806 44515: contig of 711 bp in length
* 44517 44616: gap of 100 bp
* 44617 45327: contig of 711 bp in length
* 45328 45427: gap of 100 bp
* 45428 46150: contig of 723 bp in length
* 46151 46250: gap of 100 bp
* 46251 46973: contig of 723 bp in length
* 46974 47073: gap of 100 bp
* 47074 47782: contig of 709 bp in length
* 47783 47882: gap of 100 bp
* 47883 48595: contig of 713 bp in length
* 48596 48695: gap of 100 bp
* 48696 49391: contig of 696 bp in length
* 49392 49491: gap of 100 bp
* 49492 50213: contig of 722 bp in length
* 50214 50313: gap of 100 bp
* 50314 51002: contig of 689 bp in length
* 51003 51102: gap of 100 bp
* 51103 51814: contig of 712 bp in length
* 51815 51914: gap of 100 bp
* 51915 52612: contig of 698 bp in length
* 52613 52712: gap of 100 bp
* 52713 53430: contig of 716 bp in length
* 53431 53530: gap of 100 bp
* 53531 54239: contig of 709 bp in length
* 54240 54339: gap of 100 bp
* 54340 55071: contig of 732 bp in length
* 55072 55171: gap of 100 bp
* 55172 55901: contig of 730 bp in length
* 55902 56001: gap of 100 bp
* 56002 56722: contig of 721 bp in length
* 56723 56822: gap of 100 bp
* 56823 57533: contig of 711 bp in length

Query Match      37.4%; Score 156; DB 2; Length 65608;
Best Local Similarity 87.2%; Pred. No. 2e-20;
Matches 171; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
```

12 GTGGCGGAGGACCTCCGACATCCAGGACATGAGAGAGACATGAGAGATGATGGA 71  
13521 GTGGCGGAGGAGAGTCCGACATCCAGGACATGAGAGAGAGAGAGAGATGGA 13462

QY 72 GAGCGACCGGACGACGATGCTGACGCTGATGCTGCCGCGCTTACCTACTGCG 131  
DB GAGCGACCTTCCACGAGATGATACATGATGCTGCCGCGCTTACCTACTGCG 13402

QY 132 CCAAGCCCAAGTGTCTGCTGACGCTGACGAGAGAGAGAGTGTGACAGCCCGGCTCAC 191  
DB 13401 CCAAGCCCAAGTGTGTATCAACTGAGGAGAGAGAGAGTGTGTGACAGCCCGGCTCAC 13342

Oy	192	CAACAGCGCCATGCCG	207
Db	1341	CAGCAGCGCCGTCCG	1326

RESULT 15	LOCUS	DEFINITION
AL645911	244653 bp	DNA linear RCD 09-AUG-2002
AL645911	Mouse DNA sequence from clone RP23-25M3 on chromosome 11, complete	

REFERENCE AUTHORS TITLE JOURNAL	COMMENT
1 (bases 1 to 244653) Tromans, A. Direct Submission Submitted (03-AUG-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: hummer@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk On Aug 11, 2002 this sequence version replaced gi:15151055.	

BASE COUNT	59844	a	60816	c	61968	g	62025	t
ORIGIN								

QY 1 ATGGGAGAACTGTGTCGCGCAGGAGACTCGGCACTACGGCACTGGACGAGAGACACTGTGG 60  
Db 24041 ATGGCAGAACTGTGTCGCGCAGGAGACTCGGCACTACGTCTGTGACGAGAGAGTCTGTGG 24100  
QY 61 GAGATGATGAGAGAGCCACCGCCACAGGATGTGACGCTGCATTCGCCCAAGCCGCTCACC 120

Qy	Db	Qy	Db
121	24101	121	24101
24161	24101	24161	24101
181	24221	181	24221
24221	24221	24221	24221

```
Search completed: February 17, 2004, 22:52:04
Job time : 2074.14 secs
```





```

FT      /*tag= b
FT      /note= "Encodes caspase recruitment domain (CARD).
FT      Specifically claimed in claim 2"
XX      US2002164703-A1.
XX      07-NOV-2002.
XX      19-DEC-2001; 2001US-0032159.
XX      21-DEC-2000; 2000US-257457P.
XX      (PAWL/) PAWLOWSKI K.
XX      (PAWL/) REED J C.
XX      (GODZ/) GODZIK A.
XX      Pawlowski K, Reed JC, Godzik A;
XX      WPI; 2003-288137/28.
XX      P-PSDB; ABG76062.
XX      New isolated CARD-containing nucleic acids, useful for the diagnosis
XX      and treatment of disorders with aberrant expression or activity of the
XX      CARD-containing polypeptide, such as cancer, stroke, arthritis, heart
XX      failure and AIDS.
XX      Claim 1; Fig 3; 34pp; English.
XX      The invention relates to an isolated nucleic acid molecule encoding a
XX      caspase recruitment domain (CARD) containing polypeptide. CARD containing
XX      polypeptides are involved in apoptosis (as caspase activators and caspase
XX      inhibitors), cell adhesion, inflammation and cytokine receptor
XX      signaling. The methods and compositions of the present invention are
XX      useful for the diagnosis and treatment of disorders associated with the
XX      aberrant expression or activity of the CARD containing polypeptide such
XX      as cancer, glioma, lymphoma, adenocarcinoma, sarcoma, melanoma, keloid,
XX      hamartoma, leukemia, lymphoma, keratinocyte hyperplasia, neoplasia,
XX      benign prostatic hypertrophy, inflammatory diseases such as arthritis, lupus,
XX      Sjogren's syndrome, Crohn's disease, ulcerative colitis, graft versus
XX      host disease, sepsis, abnormal cell death diseases such as stroke,
XX      myocardial infarction, heart failure, neurodegenerative diseases like
XX      Parkinson's disease and Alzheimer's disease, and HIV infection. The
XX      present sequence represents cDNA encoding the human caspase recruitment
XX      domain containing protein, CARD-12X.
XX      Sequence 417 BP; 84 A; 131 C; 128 G; 72 T; 2 other;
SQ
Query Match      99.5%; Score 415; DB 25; Length 417;
Best Local Similarity 100.0%; Pred. No. 5.4e-83;
Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Db      301 ACCCTGTCACCGGGCTGCAAGCTGATGTTGACTTCAATTAAGCGTGAAGCTCC 360
QY      361 GACTTGAACGGTTGGACAGCACTTGTAGAACTCAGACTCCTGTGAACCCAG 415
Db      361 GACTTGAACGGTTGGACAGCACTTGTAGAACTCAGACTCCTGTGAACCCAG 415

RESULT 2
AAD13448
ID      AAD13448 standard; DNA; 3417 BP.
XX
XX      AAD13448;
XX
XX      06-NOV-2001 (first entry)
XX
XX      Human caspase recruitment domain-14 (CARD-14) genomic DNA.
XX
XX      Human, caspase recruitment domain-14; CARD-14; chromosome 17;
XX      nuclear factor-kappa B; NF-kB; cell growth; cell death; cancer; therapy;
XX      autoimmune disorder; systemic lupus erythematosus; neurological disorder;
XX      Alzheimer's disease; Parkinson's disease; inflammatory disorder; anaemia;
XX      haematological disorder; myelodysplastic syndrome; myocardial infarction;
XX      stroke; immune disorder; Crohn's disease; allergic rhinitis; infection;
XX      cell signalling disorder; cytostatic; immunosuppressive; nootropic;
XX      neuroprotective; antiviral; antibacterial; ds.
XX
XX      Homo sapiens.
XX
XX      Location/Qualifiers
XX      FH      Key
XX      FT      CDS
XX      FT      1..3417
XX      FT      /*tag= a
XX      FT      /product= "Human caspase recruitment domain-14"
XX      FT      /transl_except= (pos:2098..2100, aa:Cys)
XX      FT      /note= "CDS does not include stop codon"
XX      FT      /partial
XX      FT      1..211
XX      FT      /*tag= b
XX      FT      /number= "1"
XX      FT      212..415
XX      FT      /*tag= c
XX      FT      /number= "2"
XX      FT      416..741
XX      FT      /*tag= d
XX      FT      /number= "3"
XX      FT      742..909
XX      FT      /*tag= e
XX      FT      /number= "4"
XX      FT      910..1029
XX      FT      /*tag= f
XX      FT      /number= "5"
XX      FT      1030..1209
XX      FT      /*tag= g
XX      FT      /number= "6"
XX      FT      1210..1359
XX      FT      /*tag= h
XX      FT      /number= "7"
XX      FT      1360..1476
XX      FT      /*tag= i
XX      FT      /number= "8"
XX      FT      1477..1619
XX      FT      /*tag= j
XX      FT      /number= "9"
XX      FT      1620..1714
XX      FT      /*tag= k
XX      FT      /number= "10"
XX      FT      1715..1778
XX      FT      /*tag= l
XX      FT      /number= "11"
XX      FT      1779..2001
XX      FT      /*tag= m
XX      FT      /number= "12"
XX      FT      2002..2128
XX      FT      exon

```

```

FT      /tag= n
FT      /number= "13"
FT      2129..2369
FT      /tag= o
FT      /number= "14"
FT      2370..2433
FT      /tag= p
FT      /number= "15"
FT      2434..2548
FT      /tag= q
FT      /number= "16"
FT      2549..2719
FT      /tag= r
FT      /number= "17"
FT      2720..2841
FT      /tag= s
FT      /number= "18"
FT      2842..2957
FT      /tag= t
FT      /number= "19"
FT      2958..3138
FT      /tag= u
FT      /number= "20"
FT      3139..3417
FT      /tag= v
FT      /number= "21"

PN      WO200159065-A2.
XX      16-AUG-2001.
PD      22-JAN-2001; 2001WO-US02087.
XX      09-FEB-2000; 2000US-0181159.
XX      (MILL-) MILLENNIUM PHARM INC.
XX      Bercin J;
XX      MPI: 2001-497073/54.
XX      P-PSDB; AAE07165.
DR      An isolated caspase recruitment domain polypeptide useful for
PT      regulating growth and cell death and useful for the treatment of cancer
PT      -
XX      Disclosure; Fig 2A-2C; 109pp; English.
XX      The present sequence is human caspase recruitment domain-14 (CARD-14)
CC      DNA. The CARD-14 gene is located on chromosome 17. The CARD-14 is used
CC      for the detection of modulators that modulates the ability of CARD-14 to
CC      bind to Bcl-10 and stimulates phosphorylation of Bcl-10 or activation
CC      of nuclear factor-kappa B (NF-kB). The CARD-14 is useful for regulating
CC      growth and cell death and useful for the treatment of cancer. It is
CC      also useful for the treatment of autoimmune disorders (e.g., systemic
CC      lupus erythematosus), neurological disorders (e.g., Alzheimer's and
CC      Parkinson's disease, inflammatory disorders, hematological disorders
CC      (e.g., anaemia, myelodysplastic syndromes), myocardial infarctions,
CC      strokes, immune disorders (e.g., Crohn's disease, allergic rhinitis),
CC      cell signalling disorders and certain viral and bacterial infections.
XX      Sequence 3417 BP, 739 A; 1039 C; 1102 G; 537 T; 0 other;
SQ

```

```

Query Match      99.5%; Score 415; DB 22; Length 3417;
Best Local Similarity 100.0%; Pred. No. 7.2e-83;
Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 ATGGGGGAACTGTGCGGAGGAGCTCCGCACTCAGCGCACTGAGAGAGAGCACTGTGG 60
DB      1 ATGGGGGAACTGTGCGGAGGAGCTCCGCACTCAGCGCACTGAGAGAGAGCACTGTGG 60
QY      61 GAGATGATGAGAGCCACCGCCACAGAGATGATGATGCTTCCCGAGCGGCTTACC 120

```

```

DB      61 GAGATGATGAGAGCCACCGCCACAGAGATGATGATGCTTCCCGAGCGGCTTACC 120
QY      121 CCTTACTCTGGCCAGGCGCAAGGTGCTGTGTGCGACGTGAGAGAGAGAGGTGTCACAGC 180
DB      121 CCTTACTCTGGCCAGGCGCAAGGTGCTGTGTGCGACGTGAGAGAGAGAGGTGTCACAGC 180
QY      181 CCCGGGCTACCAACAGCGCCATGCGGGGCGGGGCACTGTGTGATTTGTGAGAGACTCGA 240
DB      181 CCCGGGCTACCAACAGCGCCATGCGGGGCGGGGCACTGTGTGATTTGTGAGAGACTCGA 240
QY      241 GGGAGAGACGGGCGCCATGCGCTTCTGAGAGAGGCTGAGTCCACACCTGACGTCTAC 300
DB      241 GGGAGAGACGGGCGCCATGCGCTTCTGAGAGAGGCTGAGTCCACACCTGACGTCTAC 300
QY      301 ACCCTGTACCGGGGCTGACGCTGATGTTGACTTCAGTAACTTTAGCGGTGAGAGCTCC 360
DB      301 ACCCTGTACCGGGGCTGACGCTGATGTTGACTTCAGTAACTTTAGCGGTGAGAGCTCC 360
QY      361 GACTTTGACGGTTTGGCAGGCACTTCTAGGAACCTCAGGCTCCTGTGAACCCAG 415
DB      361 GACTTTGACGGTTTGGCAGGCACTTCTAGGAACCTCAGGCTCCTGTGAACCCAG 415

RESULT 3
AADI3447
ID      AADI3447 standard; cDNA; 3931 BP.
XX      AC
AC      AADI3447;
XX      06-NOV-2001 (first entry)
DE      Human caspase recruitment domain-14 (CARD-14) cDNA.
XX      XX
KM      Human, caspase recruitment domain-14; CARD-14; chromosome 17;
KM      nuclear factor-kappa B; NF-kB; cell growth; cell death; cancer; therapy;
KM      autoimmune disorder; systemic lupus erythematosus; neurological disorder;
KM      Alzheimer's disease; Parkinson's disease; inflammatory disorder; anaemia;
KM      haematological disorder; myelodysplastic syndrome; myocardial infarction;
KM      stroke; immune disorder; Crohn's disease; allergic rhinitis; infection;
KM      cell signalling disorder; cytostatic; immunosuppressive; neurotropic;
KM      neuroprotective; antiviral; antibacterial; ss.
XX      OS
OS      Homo sapiens.
XX      XX
FH      Key
FH      CDS
FT      Location/Qualifiers
FT      207..3221
FT      /tag= a
FT      /product= "Human caspase recruitment domain-14"
FT      /note= "Residues 207-3218 is specifically claimed
FT      as SEQ ID NO 3 in claim 9 of the specification"
XX      WO200159065-A2.
XX      16-AUG-2001.
XX      22-JAN-2001; 2001WO-US02087.
XX      09-FEB-2000; 2000US-0181159.
XX      (MILL-) MILLENNIUM PHARM INC.
XX      Bercin J;
XX      MPI: 2001-497073/54.
XX      P-PSDB; AAE07164.
DB      An isolated caspase recruitment domain polypeptide useful for
PT      regulating growth and cell death and useful for the treatment of cancer
PT      -
XX      Claim 9; Fig 1A-1E; 109pp; English.
XX      The present sequence is human caspase recruitment domain-14 (CARD-14)
CC

```



```

FT      /*tag= a
FT      /product= "CARD-10"
FT      /note= "ORF is specifically claimed"
XX
XX      MO200140468-A2.
XX
XX      07-JUN-2001.
XX
XX      01-DEC-2000; 2000WO-US32716.
XX
XX      03-DEC-1999; 99US-0168780.
XX      18-FEB-2000; 2000US-0507533.
XX      25-FEB-2000; 2000US-0513904.
XX      10-OCT-2000; 2000US-0685791.
XX
XX      (MILL-) MILLENNIUM PHARM INC.
XX
XX      Bertin J;
XX
XX      WPI; 2001-367809/38.
XX      P-PSDB; AAU01206.
XX
XX      Novel caspase recruitment domain (CARD) proteins, CARD-9, CARD-10,
XX      CARD-11, useful as targets for therapy, as immunogens, and in screening
XX      and detection assays -
XX
XX      Claim 2; Fig 10A-10C; 145bp; English.
XX
XX      The present sequence encoding for novel human caspase recruitment
XX      domain, CARD-10 is isolated from a human skin cDNA library.
XX      Also described are novel human sequences for CARD-9 and CARD-11
XX      (AAU01205, AAU01207) and rat CARD-9 (AAU01204). CARD-9, CARD-10 and
XX      CARD-11 interact with Bcl-10 which is thought to activate nuclear factor
XX      (NF)-kappaB and apoptosis. The sequences of the invention can be used for
XX      treating a disorder associated with abnormal levels of apoptosis by
XX      modulating the expression or activity of CARD-9, CARD-10, or CARD-11.
XX      They can be used for the treatment of hyperproliferative disorders
XX      (e.g. cancer), autoimmune disorders (e.g. systemic lupus erythematosus),
XX      neurological disorders (e.g. Alzheimer's disease), inflammatory disorders
XX      (e.g. Crohn's disease), and viral infection (e.g. HIV). The CARD
XX      polypeptide, polynucleotide and an antibody which selectively binds to
XX      CARD can be used in screening and detection assays (e.g. chromosomal
XX      mapping, tissue typing), predictive medicine (prognostic assays,
XX      monitoring clinical trials, and therapy (treatment and prophylaxis). The
XX      CARD polypeptide may be used to screen for drugs that bind to and/or
XX      modulate it. CARD sequences are potential targets for regulating
XX      inflammation, cancer, NF-kappaB signaling, stress-related response and
XX      apoptosis in human disease. A host cell containing a polynucleotide
XX      encoding CARD can be used to create transgenic animals.
XX
XX      Sequence 3949 BP; 724 A; 1222 C; 1319 G; 684 T; 0 other;
XX
XX      Query Match      28.2%; Score 117.8; DB 22; Length 3949;
XX      Best Local Similarity 59.3%; Pred. No. 6.2e-17;
XX      Matches 200; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

```

```

DB      322 GGGCTTCCTGAGACCCCTGGAGTCTACTACCCCGAACCTTCAAGTCTCAGGGCCA 381
OY      318 GCAGCTGATGTTGACTTCACTTAATTAGCGGTGAG 354
DB      382 GGAACCGCGCCGAGCGCTGCTCATGATCTCTGATGAG 418

RESULT 6
ABA00333
ID      ABA00333 standard; cDNA; 3949 BP.
XX
XX      ABA00333;
XX
XX      09-DEC-2002 (first entry)
XX
XX      Human CARD-10 cDNA.
XX
XX      Gene; rat; human; caspase recruitment domain; CARD-9; CARD-10;
XX      CARD-11; apoptosis; inflammation; cell growth; cell death;
XX      lymphocyte activation; cancer; melanoma; autoimmune disease;
XX      arthritis; neurological disorder; Alzheimer's disease; ss.
XX
XX      Homo sapiens.
XX
XX      Key      location/Qualifiers
XX      CDS      41..3139
XX      FT      /*tag= a
XX      FT      /product= "CARD-10"
XX
XX      MO200270652-A2.
XX      12-SEP-2002.
XX
XX      28-FEB-2002; 2002WO-US06147.
XX
XX      02-WAR-2001; 2001US-0798412.
XX
XX      (MILL-) MILLENNIUM PHARM INC.
XX
XX      Bertin J;
XX
XX      WPI; 2002-698749/75.
XX      P-PSDB; AAG79554.
XX
XX      CARD-9, CARD-10 or CARD-11 polypeptides and polynucleotides, useful for
XX      treating disorders associated with inappropriate apoptosis or
XX      lymphocyte activation, e.g. cancer -
XX
XX      Claim 5; Fig 10; 151bp; English.
XX
XX      This sequence encodes human caspase recruitment domain (CARD)-10.
XX      CARD proteins play roles in apoptotic and inflammatory signalling
XX      pathways. CARD-9, -10 and -11 participate in the network of
XX      interactions that modulate caspase activity. They are thought to be
XX      useful as modulating agents for regulating a variety of cellular
XX      processes including cell growth and cell death. CARD proteins and
XX      nucleic acids are useful for treating a disorder associated with
XX      inappropriate apoptosis or lymphocyte activation or for diagnosing
XX      subjects having or that are at risk of developing a disorder associated
XX      with aberrant CARD-9, CARD-10 or CARD-11 expression or activity, such
XX      as cancer e.g. melanoma, autoimmune disorders e.g. arthritis, or
XX      neurological disorders e.g. Alzheimer's disease.
XX
XX      Sequence 3949 BP; 724 A; 1222 C; 1319 G; 684 T; 0 other;
XX
XX      Query Match      28.2%; Score 117.8; DB 24; Length 3949;
XX      Best Local Similarity 59.3%; Pred. No. 6.2e-17;
XX      Matches 200; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

```







```

Db      357 GAAAGAGCGCTTGTGGAGAAATGTGAGTAAACCGGCAACAGCTCAGCGCTATATC
Qy      103 TCCCCAGCCGCTTACACCCCTTACCTGGCCAGGCGCAAGGTGTGTCCAGCTGACGAG
Db      417 AACCTGCGAAGTCAACCCCTTACTGTGTGTAAGTCAATTGATGAGCAGATGAA
Qy      163 GAGAGGTGTCTGCAACGCCCCGCGCTCAACAAACGCGCATGCGGCGCGCACTTGTCTG
Db      477 GATGAAGTCTTAATAGCCCTATGTCTGCTGCAAGATCAACGAGAGGCGCGCTGTG
Qy      223 GATTGCTGAAGACTCGAGGGAAGAACGGGCGCATGCGCTTCTTGGAGAGCTTGAATTC
Db      537 GACATTTCACTACATCCAGGCGCAAGGGGCTATGTGTCTTCTTGGAGAGCTTGAATTT
Qy      283 CACAACCTGACGCTTACACCGCTGTGACCGGCGCTGACGCTGATGTTGACTTCAAGTAC
Db      597 TATTACCGAAGACTGTACAACTGGTGACTGGGAAAGAGCCCACTCGGAGATTCTCAC
Qy      343 TTTAGCGGTGAG 354
Db      657 ATTGTGTGAG 668

```

## RESULT 10

ABA00334  
ID ABA00334 standard; cDNA; 4276 BP.

XX ABA00334;

DT 09-DEC-2002 (first entry)

DE Human CARD-11 cDNA.

XX Gene; rat; human; caspase recruitment domain; CARD-9; CARD-10;  
KM CARD-11; apoptosis; inflammation; cell growth; cell death;  
KM lymphocyte activation; cancer; melanoma; autoimmune disease;  
KM arthritis; neurological disorder; Alzheimer's disease; ss.

OS Homo sapiens.

XX Key Location/Qualifiers  
FH 328..3771  
FT /tag=a  
FT /product="CARD-11"

PN WO200270652-A2.

PD 12-SEP-2002.

PF 28-FEB-2002; 2002WO-US06147.

PR 02-MAR-2001; 2001US-0798412.

PA (MILL-) MILLENNIUM PHARM INC.

PI Bertin J;

XX WPI; 2002-698749/75.  
XX P-P8DB; AAG79555.

PT CARD-9, CARD-10 or CARD-11 polypeptides and polynucleotides, useful for  
PT treating disorders associated with inappropriate apoptosis or  
PT lymphocyte activation, e.g. cancer

PS Disclosure; Fig 14; 151pp; English.

CC This sequence encodes human caspase recruitment domain (CARD)-11.  
CC CARD proteins play roles in apoptotic and inflammatory signalling  
CC pathways. CARD-9, -10 and -11 participate in the network of  
CC interactions that modulate caspase activity. They are thought to be  
CC useful as modulating agents for regulating a variety of cellular  
CC processes including cell growth and cell death. CARD proteins and

CC nucleic acids are useful for treating a disorder associated with  
CC inappropriate apoptosis or lymphocyte activation or for diagnosing  
CC subjects having or that are at risk of developing a disorder associated  
CC with aberrant CARD-9, CARD-10 or CARD-11 expression or activity, such  
CC as cancer e.g. melanoma, autoimmune disorders e.g. arthritis, or  
CC neurological disorders e.g. Alzheimer's disease.

XX Sequence 4276 BP; 1045 A; 1256 C; 1320 G; 655 T; 0 other;

Qy Query Match 24.9%; Score 104; DB 24; Length 4276;

Best Local Similarity 58.3%; Pred. No. 7.2e-14; Matches 182; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

```

Qy      43 GAGAGAGACACTGTGGAGATGATGAGAGCCACCGGCAAGATCTGACGCTGCATC
Db      358 GAAAGAGACCGCTTGTGGAGATGATGAGATGTAAACCGGCAATGCTACCGCTATATC
Qy      103 TCCCCAGCCGCTTACACCCCTTACCTGGCCAGGCGCAAGGTGTGTCCAGCTGACGAG
Db      418 AACCTGCGAAGTCAACCCCTTACTGTGTGTAAGTCAATTGATGAGCAGATGAA
Qy      163 GAGAGGTGTCTGCAACGCCCCGCGCTCAACAAACGCGCATGCGGCGCGCACTTGTCTG
Db      478 GATGAAGTCTTAATAGCCCTATGTCTGCTGCAAGATCAACGAGAGGCGCGCTGTG
Qy      223 GATTGCTGAAGACTCGAGGGAAGAACGGGCGCATGCGCTTCTTGGAGAGCTTGAATTC
Db      538 GACATTTCACTACATCCAGGCGCAAGGGGCTATGTGTCTTCTTGGAGAGCTTGAATTT
Qy      283 CACAACCTGACGCTTACACCGCTGTGACCGGCGCTGACGCTGATGTTGACTTCAAGTAC
Db      598 TATTACCGAAGACTGTACAACTGGTGACTGGGAAAGAGCCCACTCGGAGATTCTCAC
Qy      343 TTTAGCGGTGAG 354
Db      658 ATTGTGTGAG 668

```

## RESULT 11

AAH08620  
ID AAH08620 standard; cDNA; 765 BP.

XX AAH08620;

DT 26-JUN-2001 (first entry)

DE Human cDNA clone (5'-primer) SEQ ID NO:5455.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

OS Homo sapiens.

PN EP1074617-A2.

PD 07-FEB-2001.

PF 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118767.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

PA (HELI-) HELIX RES INST.

XX Oca T, Isegai T, Mshikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX WPI; 2001-318749/34.

PT Primer sets for synthesizing polynucleotides, particularly the 5602  
PT full-length cDNAs defined in the specification, and for the detection

PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -

PS Claim 1; SEQ ID 5455; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH01366 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.

XX Sequence 765 BP; 160 A; 236 C; 242 G; 123 T; 4 other;

XX Query Match 22.6%; Score 94.4; DB 22; Length 765;

XX Best Local Similarity 55.5%; Pred. No. 7.7e-12;  
 XX Matches 182; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

QY 42 GGACGAGAGAGACACTGTGGAGATATGAGAGCCACCGCCACAGATCTGACGTGAT 101  
 DB 161 GAAACATGACGAGTCTGAGACGTCCTGAGGGCTTCCGGGTGAGCTCACCCTCGTAT 220  
 QY 102 CTGCCACGAGCCGCTCAACCTTACCTGTGGCCAGGCGCAAGGCTGTGACAGTGAAGA 161  
 DB 221 GACCCCTTACGACATCACTTACCTTACCTGCGAGGCAAGGCTTGAACCTGATGATGA 280  
 QY 162 GGAGGAGGTGCTGCAAGCGCCCGCTCAACAAGCGCCATGCGGGCGGCACTTCT 221  
 DB 281 GGAGGAGGTGCTGCAAGCGCCCGCTCAACAAGCGCCATGCGGGCGGCTCTCT 340  
 QY 222 GATTTGCTGAAGACTCGAGGAGAAAGACGGGCGCATTCGCTTCTGAGAGCTGAAGTT 281  
 DB 341 GGACATCTCTGACGCGACCGGCGCACAAAGGCTACCTGCGCTTCTCGAGAGCTGAGACT 400  
 QY 282 CCACAACCTCTGACGCTTACCTGTGTCACCGGGCTGACAGCTGATGTTGACTTCAAGTAA 341  
 DB 401 CTACTACCGCGAGCTGTACAGAAAGTCAACAGCAAGAGCGGCGCGCTTCTTCCAT 460  
 QY 342 CTTTACGGGTGAGAGCTCCGACTTTGAC 369  
 DB 461 GATCATCGACGCTCCGGGAGTCAAGC 488

RESULT 12

AA05387  
 ID AA05387 standard; cDNA; 2098 BP.

XX AA05387;

XX 12-SEP-2001 (first entry)

XX Human caspase recruitment domain, CARD-9 cDNA sequence.

XX Human; caspase recruitment domain; CARD-9; Bcl-10; NF-kappaB;  
 KW apoptosis; hyperproliferative disorder; autoimmune; neurological;  
 XX inflammatory disorder; viral infection; stress-related response; ss.

OS Homo sapiens.

XX Key Location/Qualifiers  
 FT 144..1754  
 FT CDS /\*tag= a  
 FT /product= "CARD-9"  
 FT /note= "ORF is specifically claimed"

XX W0200140468-A2.

XX 07-JUN-2001.

XX 01-DEC-2000; 2000NO-US32716.

XX 03-DEC-1999; 99US-0168780.

XX 18-FEB-2000; 2000US-0507533.

XX 25-FEB-2000; 2000US-0513904.

XX 10-OCT-2000; 2000US-0685791.

XX (MILL-) MILLENNIUM PHARM INC.

XX Bertin J;

XX WPI: 2001-367809/38.

XX P-PSDB; AAU01205.

XX Claim 2; Fig 5A-5B; 145pp; English.

XX The present sequence encoding for novel human caspase recruitment  
 CC domain, CARD-9 is isolated from a human megakaryocyte cDNA library.  
 CC Also described are novel human sequences for CARD-10 and CARD-11  
 CC (AAU01206, AAU01207) and rat CARD-9 (AAU01204). CARD-9, CARD-10 and  
 CC CARD-11 interact with Bcl-10 which is thought to activate nuclear factor  
 CC (NF)-kappaB and apoptosis. The sequences of the invention can be used for  
 CC treating a disorder associated with abnormal levels of apoptosis by  
 CC modulating the expression or activity of CARD-9, CARD-10, or CARD-11.  
 CC They can be used for the treatment of hyperproliferative disorders  
 CC (e.g. cancer), autoimmune disorders (e.g. systemic lupus erythematosus),  
 CC neurological disorders (e.g. Alzheimer's disease), inflammatory disorders  
 CC (e.g. Crohn's disease), and viral infection (e.g. HIV). The CARD  
 CC polypeptide, polynucleotide and an antibody which selectively binds to  
 CC CARD can be used in screening and detection assays (e.g. chromosomal  
 CC mapping, tissue typing), predictive medicine (prognostic assays,  
 CC monitoring clinical trials, and therapy (treatment and prophylaxis)). The  
 CC CARD polypeptide may be used to screen for drugs that bind to and/or  
 CC modulate it. CARD sequences are potential targets for regulating  
 CC inflammation, cancer, NF-kappaB signalling, stress-related response and  
 CC apoptosis in human disease. A host cell containing a polynucleotide  
 CC encoding CARD can be used to create transgenic animals.

XX Sequence 2098 BP; 455 A; 644 C; 702 G; 297 T; 0 other;

XX Query Match 22.6%; Score 94.4; DB 22; Length 2098;

XX Best Local Similarity 55.5%; Pred. No. 8.8e-12;  
 XX Matches 182; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

QY 42 GGACGAGAGAGACACTGTGGAGATATGAGAGCCACCGCCACAGATCTGACGTGAT 101

DB 158 GAAACATGACGAGTCTGAGACGTCCTGAGGGCTTCCGGGTGAGCTCACCCTCGTAT 217

QY 102 CTGCCACGAGCCGCTCAACCTTACCTGTGGCCAGGCGCAAGGCTGTGACGCTGAGCA 161

DB 218 GACCCCTTACGACATCACTTACCTGCGGCAAGTGTGAACCTGATGATGA 277

QY 162 GGAAGAGGTGCTGCAAGCGCCCGGCTCAACAAGCGCCATGCGGGCGGGAAGCTTGTCT 221

DB 278 GGAGCAGAGTCTGACGAGACCGCAACTGCTGATCCGCAACGGAAGTGGTGTGCTCT 337

QY 222 GATTTGCTGAAGACTCGAGGAGAAAGACGGGCGCATGCGCTTCTGAGAGCTGAAGTT 281

DB 338 GGACATCTGTGAGGAGCCGAGCCACAGGAGCTACGTGGCTCTCTCCAGAGCCCTGAGCT 397  
 QY 282 CCACACCTCTGACCTCTACACCCCTGTGACCGGCTGAGCTATGTTGACTTCACTTA 341  
 DB 398 CTACTACCCCGAGCTGTACAGAGGTCACAGGAGAGGACCGCGCTCTTCTCCAT 457  
 QY 342 CTTAGCGGTGAGAGCTCCGACTTTGAC 369  
 DB 458 GATCATCGACCGCTCCGGGAGTCAGGC 485

## RESULT 13

ABA00332  
 ID ABA00332 standard; cDNA, 2098 BP.

AC ABA00332;

DT 09-DEC-2002 (first entry)

DE Human CARD-9 cDNA.

Gene; rat; human; caspase recruitment domain; CARD-9; CARD-10;  
 CARD-11; apoptosis; inflammation; cell growth; cell death;  
 lymphocyte activation; cancer; melanoma; autoimmune disease;  
 arthritis; neurological disorder; Alzheimer's disease; ss.

OS Homo sapiens.

Key Location/Qualifiers

FT CDS 144..1754

FT /\*tag= a

FT /product= "CARD-9"

PN WO200270652-A2.

PD 12-SEP-2002.

PF 28-FEB-2002; 2002WO-US06147.

PR 02-MAR-2001; 2001US-0798412.

PA (MILL-) MILLENNIUM PHARM INC.

PI Bertin J;

DR WPI: 2002-698749/75.

DR P-PSDB; AAG79553.

CARD-9, CARD-10 or CARD-11 polypeptides and polynucleotides, useful for  
 treating disorders associated with inappropriate apoptosis or  
 lymphocyte activation, e.g. cancer -

PS Disclosure; Fig 5; 151pp; English.

This sequence encodes human caspase recruitment domain (CARD)-9.  
 CARD proteins play roles in apoptotic and inflammatory signalling  
 pathways. CARD-9, -10 and -11 participate in the network of  
 interactions that modulate caspase activity. They are thought to be  
 useful as modulating agents for regulating a variety of cellular  
 processes including cell growth and cell death. CARD proteins and  
 nucleic acids are useful for treating a disorder associated with  
 inappropriate apoptosis or lymphocyte activation or for diagnosing  
 subjects having or that are at risk of developing a disorder associated  
 with aberrant CARD-9, CARD-10 or CARD-11 expression or activity, such  
 as cancer e.g. melanoma, autoimmune disorders e.g. arthritis, or  
 neurological disorders e.g. Alzheimer's disease.

Sequence 2098 BP, 455 A, 644 C, 702 G, 297 T, 0 other;

Query Match 22.6%; Score 94.4; DB 24; Length 2098;

Best Local Similarity 55.5%; Pred. No. 8-8e-12;  
 Matches 182; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

QY 42 GGACGAGGAGACCTGTGGAGATGTGAGAGCCACGAGCAAGATGTACGCTGAT 101  
 DB 158 GAACGATGACGAGTGTGAGACGTCCTGAGGCTTCCGGTGAAGCTTACCTCGAT 217  
 QY 102 CTGCCCCAGCCGCTCACCCCTTACCTGCGCCAGGCAAGTGTGTGCACTGAGCA 161  
 DB 218 CAGCCCTTACCGCATACACCTTACTGTGGGAGGCAAGTGTCTGAACCTGATGATGA 277  
 QY 162 GAGAGAGTGTGTGACAGGCCCCCGGCTACCAACGCGCATGTGCGGCGGCACTTGT 221  
 DB 278 GAGCAGGTGTCTCAAGGACCCCACTGGTCACTCGCAACGGAAGTGGGTGTCTCCT 337  
 QY 222 GGAATTGCTGAGACCTCGAGGGAAGACGGGCACTGCTTCTGTGAGAGCTTGAAT 281  
 DB 338 GGACATCTGTGACCGGACCGGCAAGAGGCTACGTGGGCTTCTCGAGAGCTGAGCT 397  
 QY 282 CCACACCTGTGACCTGTACACCCCTGTGACCGGCTGACCTGTGATGTTGACTTA 341  
 DB 398 CTACTACCCCGAGCTGTACAGAGGTCACAGGAGAGGACCGCGCTCTTCTCCAT 457  
 QY 342 CTTAGCGGTGAGAGCTCCGACTTTGAC 369  
 DB 458 GATCATCGACCGCTCCGGGAGTCAGGC 485

## RESULT 14

AAH18321  
 ID AAH18321 standard; cDNA, 2176 BP.

AC AAH18321;

DT 26-JUN-2001 (first entry)

DE Human cDNA sequence SEQ ID NO:18327.

Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

OS Homo sapiens.

PN EP1074617-A2.

PD 07-FEB-2001.

PF 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JUN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

PA (HELI-) HELIX RES INST.

PI Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

DR WPI: 2001-318749/34.

PR Primer sets for synthesizing polynucleotides, particularly the 5602

PT full-length cDNAs defined in the specification, and for the detection

PT and/or diagnosis of the abnormality of the proteins encoded by the

PT full-length cDNAs -

PS Claim 8; SEQ ID 18327, 2537pp + CD ROM; English.

The present invention describes primer sets for synthesizing 5602

full-length cDNAs defined in the specification. Where a primer set

comprises: (a) an oligo-dT primer and an oligonucleotide complementary

to the complementary strand of a polynucleotide which comprises one of

the 5602 nucleotide sequences defined in the specification, where the

oligonucleotide comprises at least 15 nucleotides; or (b) a combination

of an oligonucleotide comprising a sequence complementary to the



Qy	282	CCACACCCCTGACGCTTACACCCCTGTCACCGGCTGACGCTGATGTTGACTTCAGTAA	341
Db	401	CTACTACCCGCACTGTTCANAGAAAGTTCACAGGCAAGAGCCGAGCCGCGCTTCTCCAT	460
Qy	342	CTTAGCGGTGAGAGCTCCGACTTTGAC	369
Db	461	GATCATGACGCGCTCCGGGAGTCAGGC	488

Search completed: February 17, 2004, 21:54:30  
 Job time : 243.091 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 17, 2004, 21:44:36 ; Search time 2341.94 Seconds  
(Without alignments)  
4327.596 Million cell updates/sec

Title: US-10-032-159A-15

Perfect score: 417  
Sequence: 1 atg999ggaactgtgcgcag.....ggtctctgttaaccacagm 417

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 1215238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

EST:  
1: em\_estbta:\*  
2: em\_estbhm:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hci:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hci:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrt1:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	351	84.2	931	13	BU956261 AGENCOURT
2	278	66.7	549	10	BF079819 230513 MA
3	264.6	63.5	1600	11	AK086176 Mus muscu
4	219.6	52.7	806	10	BF100075 601752245

5	153.2	36.7	492	12	BQ031835
6	118.8	28.5	1933	9	AU090784
7	117.8	28.2	1030	14	CA489572
8	115	27.6	392	10	BF717357
9	112	26.9	618	14	CA589549
10	111.6	26.8	1141	10	BE867544
11	110.2	26.4	732	14	CD349577
12	110.2	26.4	846	10	BE287025
13	106.6	25.6	645	9	AJ455484
14	106.6	25.6	769	9	AJ453294
15	105.8	25.4	584	9	AJ455489
16	102	24.5	655	12	BF649379
17	101.4	24.3	494	10	BF601607
18	94.4	22.6	722	12	BF905841
19	94.4	22.6	738	12	BF909365
20	94.4	22.6	765	9	AU142752
21	94.4	22.6	775	12	BF907340
22	94.2	22.6	739	13	BF975798
23	92.8	22.3	498	12	BF524042
24	92.8	22.3	819	12	BF524042
25	92.8	22.3	829	10	BF307399
26	92.8	22.3	1104	13	BF358485
27	92.8	22.3	1132	12	BF922630
28	91.8	22.0	455	12	BF821044
29	91.8	22.0	479	12	BF148244
30	91.4	21.9	451	14	CB786793
31	90.4	21.7	755	12	BF1912159
32	80.4	19.3	768	12	BF137662
33	79	18.9	678	12	BF148452
34	74	17.7	425	14	CB696535
35	73.2	17.6	677	10	BF474019
36	72.4	17.4	601	10	BE291786
37	70.8	17.0	1024	14	BF702884
38	70.8	17.0	1075	11	AK002346
39	67.8	16.3	945	29	CNS021UH
40	66.8	16.0	438	14	CA389497
41	62.2	14.9	465	29	CNS04LTM
42	60.6	14.5	238	29	CNS04M7L
43	59.6	14.3	736	29	CNS037FO
44	59.2	14.2	716	29	CNS02NN7
45	56	13.4	566	10	BE040888

## ALIGNMENTS

RESULT 1  
BU956261 931 bp mRNA  
LOCUS AGENCOURT 10613482 NIH MGC 107 Homo sapiens CDNA clone  
DEFINITION IMAGE:6725809 5', mRNA sequence.

ACCESSION BU956261.1 GI:24185833  
VERSION BU956261  
KEYWORDS EST.

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 931)

National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished

CONTACT: Robert Strausberg, Ph.D.  
Email: cga@nci.nih.gov

CDNA Library Preparation: Ruben Laboratory  
Tissue Procurement: ARCC

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov

plate: LICM053 row: 1 column: 16  
High quality sequence stop: 594.

## FEATURES

Location/Qualifiers  
1. 931

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6729809"  
/tissue\_type="adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIM\_MGC\_107"  
/note="Organ: Breast; Vector: pOTB7, Site 1: EcoRI; Site 2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIM\_MGC library."

BASE COUNT 201 a 294 c 295 g 140 t 1 others  
ORIGIN

Query Match 84.2%; Score 351; DB 13; Length 931;  
Best Local Similarity 100.0%; Pred. No. 2.6e-66;  
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGGAATGTGCGCGCAGGGACTCGCACTACGCGCACTGAGAGAGACACTGTGG 60  
DB 184 ATGGGGGAATGTGCGCGCAGGGACTCGCACTACGCGCACTGAGAGAGACACTGTGG 243  
QY 61 GAGATGATGAGAGACCAAGGATGTAAGTGTGATGCGCCCAAGCGGCTCAGC 120  
DB 244 GAGATGATGAGAGACCAAGGATGTAAGTGTGATGCGCCCAAGCGGCTCAGC 303  
QY 121 CCTTACTGCGCAGGCGCAAGTGTGTGTGACAGTGGACGAGAGAGAGTGTGCAAGC 180  
DB 304 CCTTACTGCGCAGGCGCAAGTGTGTGTGACAGTGGACGAGAGAGAGTGTGCAAGC 363  
QY 181 CCGCGGCTCACCAACAGCGCATGCGGGCGGCGCACTTGTGATTTCTGAAGACTGCA 240  
DB 364 CCGCGGCTCACCAACAGCGCATGCGGGCGGCGCACTTGTGATTTCTGAAGACTGCA 423  
QY 241 GGGAGAAAGGGGGCGATCGCTTCTGTGAGAGAGCTGAAGTTCACAACTTGAAGCTTAC 300  
DB 424 GGGAGAAAGGGGGCGATCGCTTCTGTGAGAGAGCTGAAGTTCACAACTTGAAGCTTAC 483  
QY 301 ACCCTGTACCGGGGCTGACGCTGATGTGACTGACTTAACCTTACCGGT 351  
DB 484 ACCCTGTACCGGGGCTGACGCTGATGTGACTTCACTTAACCTTACCGGT 534

RESULT 2 549 bp mRNA linear EST 18-OCT-2000  
BF079819  
DEFINITION 230513 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.  
ACCESSION BF079819  
VERSION BF079819.1 GI:10873649  
KEYWORDS EST.  
SOURCE Sus scrofa (pig)  
ORGANISM Sus scrofa

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
1 (bases 1 to 549)  
Fahrenkrug, S.C., Smith, T.P.L., Freking, B.A., Cho, J., White, J.,  
Valter, J., Wise, T., Kotter, G.A., Petter, G., Sultana, R., Quackenbush,  
'J. and Keele, J.W.  
Porcine gene discovery by normalized cDNA-library sequencing and  
EST cluster assembly

TITLE Mamm. Genome 13 (8), 475-478 (2002)  
JOURNAL 22213789  
MEDLINE 12226715  
COMMENT Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366

## FEATURES

source  
Location/Qualifiers  
1. 549  
/organism="Sus scrofa"  
/mol\_type="mRNA"  
/db\_xref="taxon:9823"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/clone\_lib="MARC 2P1G"  
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI; Library made from pooled tissue from testis, ovary, endometrium, hypothalamus, pituitary, and placenta."

BASE COUNT 102 a 169 c 188 g 90 t  
ORIGIN

Query Match 66.7%; Score 278; DB 10; Length 549;  
Best Local Similarity 87.1%; Pred. No. 1.6e-50;  
Matches 305; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 1 ATGGGGGAATGTGCGCGCAGGGACTCGCACTACGCGCACTGAGAGAGACACTGTGG 60  
DB 71 ATGGGGGAATGTGCGCGCAGGGACTCGCACTACGCGCACTGAGAGAGAGACTGTGG 130  
QY 61 GAGATGATGAGAGACCAAGGATGTAAGTGTGATGCGCCCAAGCGGCTCAGC 120  
DB 131 GAGATGATGAGAGACCAAGGATGTAAGTGTGATGCGCCCAAGCGGCTCAGC 190  
QY 121 CCTTACTGCGCAGGCGCAAGTGTGTGTGACAGTGGACGAGAGAGAGTGTGCAAGC 180  
DB 191 CCTTACTGCGCAGGCGCAAGTGTGTGTGACAGTGGACGAGAGAGAGTGTGCAAGC 250  
QY 181 CCGCGGCTCACCAACAGCGCATGCGGGCGGCGCACTTGTGATTTCTGAAGACTGCA 240  
DB 251 CCGAGGTTACCAACAGCGCATGAGATTGGGCACTTGTGATTTCTGAAGACTGCA 310  
QY 241 GGGAGAAAGGGGGCGATCGCTTCTGTGAGAGAGCTGAAGTTCACAACTTGAAGCTTAC 300  
DB 311 GGGAGAAAGGGGGCGATCGCTTCTGTGAGAGAGCTGAAGTTCACAACTTGAAGCTTAC 370  
QY 301 ACCCTGTACCGGGGCTGACGCTGATGTGACTGACTTAACCTTACCGGT 350  
DB 371 ACCCTGTACCGGGGCTGACGCTGATGTGACTTCACTTAACCTTACCGGT 420

RESULT 3 1600 bp mRNA linear HTC 05-DEC-2002  
AK086176  
DEFINITION Mus musculus 15 days embryo head cDNA, RIKEN full-length enriched library, clone:DB30010119 product:CDSPASE RECRUITMENT DOMAIN PROTEIN 14 (BDL0-INTERACTING MAGUK PROTEIN 2) (BIMF2) [Mus musculus], full insert sequence.  
ACCESSION AK086176  
VERSION AK086176.1 GI:26103277  
KEYWORDS HTC, CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
JOURNAL 99279253  
MEDLINE 10349636  
PUBMED



Tissue Procurement: Gilbert Smith, Ph.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILN)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LILN at:  
<http://image.llnl.gov>  
 Plate: LILN9174 row: j column: 12  
 High quality sequence stop: 668.  
 Location/Qualifiers

## FEATURES

source

1. 806  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:3979883"  
 /issue\_type="tumor, biopsy sample"  
 /dev\_stage="10 months, virgin"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI CGAP Mam1"  
 /note="Organ: mammary; Vector: pCMV-Sport6; Site 1: SalI;  
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
 Library constructed by Life Technologies. Investigator  
 providing samples: Gilbert Smith, NIH"

BASE COUNT  
 188 a 214 c 242 g 162 t

Query Match 52.7%; Score 219.6; DB 10; Length 806;  
 Best Local Similarity 83.6%; Pred. No. 8.7e-38;  
 Matches 296; Conservative 0; Mismatches 54; Indels 4; Gaps 4;

QY 1 ATGGGGGAAGTGTGGCGGAGGAGCTCCGCACTCAACGCGACTGACGAGAGAGCACTGTGG 60  
 DB 304 ATGGCGAAGTGTGGCGGAGGAGCTCCGCACTCAACGCGACTGACGAGAGAGTGTCTGG 363  
 QY 61 GAGATGATGAGAGCCAGCCGCAAGATGCTGACGCTGATCTGCCCCGCGCTCACC 120  
 DB 364 GATATGCTGAGAGTCAAGATGACAGATGTCAGAGATCTGCCCTGACCGGCTCACCC 423  
 QY 121 CCTTACCTGGCGGAGGAGGAGTGTGCTGCACTGACGAGAGAGAGTGTGCTGACAGC 180  
 DB 424 CCTTACCTGGCGGAGGAGGAGTGTGCTGCACTGACGAGAGAGAGTGTGCTGACAGC 483  
 QY 181 CCCCCGCTACCAACAGCGCCATGCGGCGCGGCACTTCTGATTTGCTGAAGATCGA 240  
 DB 484 TCCCGTTTACCAACAGCGCTTCTGGAAGCGCTGAAGTTCACAGAGCCCTGATGTCTGA 542  
 QY 241 GGGAGAGACGGGGCCATCGCTTCTGGAAGCGCTGAAGTTCACAA-CCCTGACGCTTA 299  
 DB 543 GGGAGAGAGTGGAGCCATGCTTCTGGAAGCGCTGAAGTTCACAGAGCCCTGATGTCTGA 602  
 QY 300 CACCCGTGCTAC-CGGGCTGACAGCTGA-TGTTGATTTCACTTAATTGCGCGT 351  
 DB 603 CACCCGTGCTACCGGGCTGACAGCTGACATTTGACATTTGACACTTTTCAAGCGGT 656

RESULT 5  
 B0031835/c 492 bp mRNA linear EST 27-MAR-2002  
 LOCUS  
 DEFINITION  
 UI-1-CFO-aoh-c-03-0-UI.s1 NCI CGAP P171 Mus musculus CDNA clone  
 B0031835  
 ACCESSION  
 VERSION  
 B0031835.1 GI:19767114  
 KEYWORDS  
 EST.  
 SOURCE  
 Mus musculus (house mouse)  
 ORGANISM  
 Mus musculus  
 BUKARYOTA; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 492)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncigap>.  
 Tumor Gene Index  
 JOURNAL  
 Unpublished

## COMMENT

Contact: Robert Strausberg, Ph.D.  
 Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
 Tissue Procurement: Dr. Minoru Ko, Dr. Janet Rosent, Dr. Janet  
 Rosent  
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be obtained  
 from Dr. M. Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
 The following repetitive elements were found in this CDNA  
 sequence: 1-49, >AT rich<low\_complexity (matched complement)  
 Seq primer: M13 FORWARD  
 POLYA=Yes.

## FEATURES

source

Location/Qualifiers  
 1. 492  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UI-1-CFO-aoh-c-03-0-UI"  
 /issue\_type="trophoblast"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="NCI CGAP P171"  
 /note="Organ: placenta; Vector: pT73-Pac (Pharmacia) with  
 a modified polylinker; Site 1: Bcl I; Site 2: Not I;  
 NCI CGAP P171 is a subcloned CDNA library constructed  
 according to Bonaldo, Lennon and Soares, Genome Research,  
 6:791-806, 1996. First strand CDNA synthesis was primed  
 with an oligo-dT primer containing a Not I site. Double  
 stranded CDNA was ligated to an Ecor I adaptor, digested  
 with Not I, and cloned directionally into pT73-Pac  
 vector. The oligonucleotide used to prime the synthesis of  
 first-strand CDNA contains a library tag sequence that is  
 located between the Not I site and the (dT)18 tail. The  
 sequence tags for this library are GGTGG, ATCAT, GGGTG,  
 GGTGG. For additional information, contact: Bento Soares,  
 bento-soares@uiowa.edu  
 TAG LIB=UI-1-CFO  
 TAG TISSUE=placenta mouse 11.5-12.5 days  
 TAG\_SEQ=ATCAT"

BASE COUNT  
 135 a 113 c 129 g 115 t

Query Match 36.7%; Score 153.2; DB 12; Length 492;  
 Best Local Similarity 79.1%; Pred. No. 2e-23;  
 Matches 182; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 1 ATGGGGGAAGTGTGGCGGAGGAGCTCCGCACTCAACGCGACTGACGAGAGAGCACTGTGG 60  
 DB 430 ATGGCGAAGTGTGGCGGAGGAGCTCCGCACTCAACGCGACTGACGAGAGAGTGTCTGG 371  
 QY 61 GAGATGATGAGAGCCAGCCGCAAGATGCTGACGCTGATCTGCCCCGCGCTCACC 120  
 DB 370 GATATGCTGAGAGTCAAGATGACAGATTTGTGAGAGCATCTGCCCTACCGGCTCAC 311  
 QY 121 CCTTACCTGGCGGAGGAGGAGTGTGCTGCACTGACGAGAGAGAGTGTGCTGACAGC 180  
 DB 310 CCTTACCTGGCGGAGGAGGAGTGTGCTGCACTGACGAGAGAGAGTGTGCTGACAGC 251  
 QY 181 CCCCCGCTACCAACAGCGCCATGCGGCGCGGCACTTCTGATTTGCT 230  
 DB 250 TCCCGTTTACCAACAGCGCCATGAGAGTGTGACTCTGTGAGATTTCTTCT 201

RESULT 6  
 A0090784 1033 bp mRNA linear EST 23-JAN-2001  
 LOCUS  
 DEFINITION  
 A0090784 lambda ZAPIT-Com A stimulated leukocytes Paralicthys  
 olivaceus CDNA clone JFconal180F forward, mRNA sequence.  
 ACCESSION  
 VERSION  
 A0090784.1 GI:12390824  
 KEYWORDS  
 EST.  
 SOURCE  
 Paralicthys olivaceus (bastard halibut)

## ORGANISM

Paralichthys olivaceus  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorphia; Acanthopterygii; Percomorphia; Pleuronectiformes;  
Pleuronectoidae; Paralichthyidae; Paralichthys.

## REFERENCE

1. (bases 1 to 1033)

## AUTHORS

Nam, B., Hirono, T. and Aoki, T.

## TITLE

Identification of expressed genes from Japanese flounder

## JOURNAL

(Paralichthys olivaceus) leukocytes stimulated with Con A/FMA

## COMMENT

Unpublished

## FEATURES

Location/Qualifiers  
1..1033  
/organism="Paralichthys olivaceus"  
/mol\_type="mRNA"  
/db\_xref="taxon:8255"  
/clone="JFconal180F"  
/cell\_type="leukocytes"  
/dev\_stage="adult"  
/clone\_lib="lambda ZAPII-Con A stimulated leukocytes"

BASE COUNT 311 a 234 c 299 g 189 t

## ORIGIN

Query Match 28.5%; Score 118.8; DB 9; Length 1033;  
Best Local Similarity 63.1%; Pred. No. 7.3e-16;  
Matches 183; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

34 ACGGCACTGACGAGAGACACTGGGAGATGATGAGGCCACGCCCAAGATGTA 93  
162 ACGCTGATGATGATGAGAGAGCTGTGGAGACGTTGGAGACCAACGCTGAACTAGT 221  
94 CGCTGCACTCTCCCGAGCGCTCAACCCCTACCTGCGCCAGCCAAAGTGTGTCAG 153  
222 TGCTTCATCTCCCGAGCAAACTCAACCCCTACCTGCGCCAGCCAAAGTGTGTCAG 281  
154 CTGACGAGAGAGAGAGTGTGTCACAGCCCGGCTCAACAGCGCCATGCGGGCGGG 213  
282 CAGATGAGAGAGAGATTCCTCACTCCCTGCTGCTGATCCAAAGCAACGAGC 341  
214 CACTTGTGATTTGCTGAGAGCTCGAGGAGAAAGGAGCCATCGCTTCTGAGAGC 273  
342 CGCTGCTGCACTCTTCACACTAAAGGAGGAGGATTCATAGGCTTTCTGAGAGT 401  
274 CTGAAGTCCCAACCCCTGAGCTACACCTGATCCACCGGGCTGACGCC 323  
402 CTGAGATTACTACCTGAGCTGACAGCTGATCAAGGAGGATCC 451

RESULT 7  
CA489572 920 bp mRNA linear EST 14-NOV-2002

LOCUS CA489572  
DEFINITION AGENCOURT\_10810481 MAPct Homo sapiens cDNA clone IMAGE:6722049 5',  
mRNA sequence.

ACCESSION CA489572  
VERSION CA489572.1 GI:24952363  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

1. (bases 1 to 920)

## AUTHORS

NIH-MGC http://mgc.ncl.nih.gov/.

## TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

## JOURNAL

Unpublished

## COMMENT

Contact: Robert Strausberg, Ph.D.

## ORGANISM

Human  
Tissue Procurement: Kristi A. Egland, Ira Pastan  
cDNA Library Preparation: Invitrogen Corp  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

## FEATURES

DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLNL4284 row: c column: 09  
High quality sequence stop: 637.  
Location/Qualifiers  
1..920

source  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6722049"  
/cell\_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, HTERT-HME1  
/lab\_host="EMD110B"  
/clone\_lib="MAPct"  
/note="Vector: pCMV-SPORT6; Site\_1: EcoRV; Site\_2: Not I;  
Subcloned with brain, liver, lung, kidney and muscle.  
Directionally cloned. Priming method: oligo-dT. Average  
insert size: 1800 bp. Library amplification: 26,000 fold.  
Kristi A. Egland, James J. Vincent, Robert Strausberg,  
Bungkok Lee & Ira Pastan: Discovery of new breast  
cancer genes encoding membrane and secreted proteins.  
Manuscript submitted."

BASE COUNT 186 a 265 c 345 g 124 t

## ORIGIN

Query Match 28.2%; Score 117.8; DB 14; Length 920;  
Best Local Similarity 61.0%; Pred. No. 1.2e-15;  
Matches 191; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

42 GAGAGAGAGACACTGTGGAGATGATGAGAGCCACGCCCAAGATGTGCTGCT 101  
6 GAGAGAGAGAGCGCTGTGGAGAGATGAGAGCGCTCCGCTGCTGCGGCTCT 65  
102 CTGCCCCAGCGCTGACCCCTACCTGCGCAGGCGCAAGTGTGTCAGCTGAGACA 161  
66 GAACCGGCGCAAGCTACCGCTGATGTCGCGCAGTCCGAGTCAAGAGAGAGACA 125  
162 GAGAGAGTGTGTCACAGCCCGGCTCAACAGCGCATGTCGGGCGGCACTTGCT 221  
126 GAGAGAGTGTGTCAGACCTACCGCTTCCGTCGCGCGTCAACCGGCGGCTGAT 185  
222 GGAATTGCTGAGAGCTCGAGGAGAAAGGAGCGCTTCTGTCGAGAGCTGAGATT 281  
186 GGCATCTTGCGCTGCGTGGCAGAGGAGGCTATGAGGCTTCTGAGAGCCCTGAGATT 245  
282 CCAACACCTGAGCTGACACCTGCTACCGGCTGTCAGGCTGAGCTGATGATTCAGTAA 341  
246 CTACTACCCCGAACAATTCACCTGCTCAAGGAGCAAGAACCGGCGGCTGCTCAT 305  
342 CTTTACGGGTGAG 354  
306 GATCTTCGATGAG 318

RESULT 8  
BG717357 392 bp mRNA linear EST 08-MAY-2001

LOCUS BG717357  
DEFINITION 602689691P1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:4821911 5',  
mRNA sequence.

ACCESSION BG717357  
VERSION BG717357.1 GI:13996544  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

1. (bases 1 to 392)

## AUTHORS

NIH-MGC http://mgc.ncl.nih.gov/.

## TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

## JOURNAL

Unpublished

## COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgaabs-r@mail.nih.gov  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki  
 Toshiyuki and Piero Carninci (RIKEN)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: LRAM10729 row: f column: 24  
 High quality sequence stop: 392.  
 Location/Qualifiers

FEATURES  
 source  
 1. 392  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4821911"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH MGC 97"  
 /note="Organ: testis; Vector: pBluescriptR (modified  
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gcgcag  
 ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTN-3',  
 size-selected for average insert size 2.2 kb and  
 normalized to ROT 5. This is a primary library enriched  
 for full-length clones and constructed using the  
 Cap-trapper method (Carninci, in preparation). Library  
 constructed by M. Brownstein (NIH/NHGRI, National  
 Institutes of Health). Note: this is a NIH MGC Library."  
 BASE COUNT 65 a 122 c 154 g 51 t  
 ORIGIN

Query Match 27.6%; Score 115; DB 10; Length 392;  
 Best Local Similarity 61.5%; Pred. No. 3.7e-15;  
 Matches 184; Conservative 0; Mismatches 115; Indels 0; Gaps 0;  
 18 CAGGAGACTCCGCACTCAGCGCACTGACGAGAGACACCTGTGGAGATGATGAGAGCA 77  
 Db 93 CGGGGGCGGCTCGGGGCTGAGGCGAGAGAGAGCGGCTGTGGAGCGAATGAGGGCGT 152  
 QY 78 CCGGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 137  
 Db 153 CCGGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 212  
 QY 138 CAGGAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 197  
 Db 213 CCGGCTATGAG 272  
 QY 198 CGCGATCGGGCGGGCGGCACTTGGCTGATGATGATGATGATGATGATGATGATGAT 257  
 Db 273 CGTACACCGGACCGGGCGCTGATGATGATGATGATGATGATGATGATGATGATGAT 332  
 QY 258 CGCCTTCTGAG 316  
 Db 333 GCGCTTCTGAG 391

RESULT 9  
 CAS89549 618 bp mRNA linear EST 19-NOV-2002  
 LOCUS hab35e06.y1 Fugu UT6 adult gut Takifugu rubripes cDNA clone  
 DEFINITION IMAGE:6350674 5', mRNA sequence.  
 ACCESSION CAS89549  
 VERSION CAS89549.1 GI:25134127  
 KEYWORDS EST.  
 SOURCE Takifugu rubripes (Fugu rubripes)  
 ORGANISM Takifugu rubripes  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 Acanthomorphi; Acanthopterygii; Percomorpha; Tetraodontiformes;  
 Tetraodontidae; Tetraodontidae; Takifugu.  
 REFERENCE 1 (bases 1 to 618)  
 AUTHORS Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy  
 S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood

TITLE  
 JOURNAL  
 COMMENT  
 Unpublished  
 Other ESTs: hab35e06.x1  
 Contact: Stephen L. Johnson  
 Washington University School of Medicine  
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: zbratishewatson.wustl.edu

K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,  
 Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schunk,R., Ritter,E.,  
 Kohn,S., Shih,T., Jackson,Y., Cardenas,X., McMan,R., Waterston,R.  
 and Wilson,R.  
 Washu Zebrafish EST Project 1998  
 Library materials provided by G. Elgar (UK MRC HGP-RC) Library  
 constructed by Drs. K. Kawakami, M. Sasaki, S. Sugano, K. Kikuchi  
 and S. Matabe (University of Tokyo, Institute of Medical Science  
 and Laboratory of Aquatic Molecular Biology and Biotechnology) DNA  
 Sequencing by: Washington University Genome Sequencing Center Clone  
 distribution: Fugu clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LNL, send email to:  
 info@image.llnl.gov  
 Seq primer: T3 RT from Amersham  
 High quality sequence stop: 480.  
 Location/Qualifiers

FEATURES  
 source  
 1. 618  
 /organism="Takifugu rubripes"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:31033"  
 /clone="IMAGE:6350674"  
 /tissue\_type="gut"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="Fugu UT6 adult gut"  
 /note="Vector: pBluescript-FL; Site 1: PflM I (CCANNNTTGG  
 ); Site 2: PflM I (CCANNNTTGG); BamHI-SmaI sites were  
 converted to BamHI-PflM-SmaI-PflM sites (SmaI is  
 destroyed). Other part of the vector is untouched. The  
 cDNA is inserted between two PflM sites in T3 (5') to T7  
 (3') direction. Library materials provided by Drs. K.  
 MRC HGP-RC) and constructed and donated by Drs. K.  
 Kawakami, M. Sasaki, S. Sugano, K. Kikuchi and S. Matabe  
 (University of Tokyo, Institute of Medical Science and  
 Laboratory of Aquatic Molecular Biology and Biotechnology  
 )."  
 BASE COUNT 176 a 148 c 170 g 124 t  
 ORIGIN

Query Match 26.9%; Score 112; DB 14; Length 618;  
 Best Local Similarity 61.8%; Pred. No. 1.9e-14;  
 Matches 178; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 36 GCGACTGAG 95  
 Db 202 GGAAGATGATGAG 261  
 QY 96 CTGATCTGCCCCAGCCGCTTCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 155  
 Db 262 CATCATCTCCCCAGGAG 321  
 QY 156 GGAG 215  
 Db 322 GGAG 381  
 QY 216 CTGCTGATTTGCTGAG 275  
 Db 382 TCTACTTGATATCTTCCACATTAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 441  
 QY 276 GAAGTTCCAAACCTGAGCTTCAACCTTGTCAACCGGCTGCAAGCC 323  
 Db 442 GGAGTTTACTACCCCGAGACTTTCAGAGCTGTCACTGGAGATGAAC 489

RESULT 10



BE867544 1141 bp mRNA linear EST 20-OCT-2000  
 LOCUS 601443043P1 NIH\_MGC\_65 Homo sapiens CDNA clone IMAGE:3847282 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BE867544  
 VERSION BE867544.1 GI:10316320  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 1141)  
 AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
 Plate: LLM9561 row: e column: 11  
 High quality sequence start: 4  
 High quality sequence stop: 269.  
 Location/Qualifiers  
 1..1141  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3847282"  
 /tissue\_type="adenocarcinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_1ib="NIH MGC 65"  
 /note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI;  
 Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 1.8 Kb. Library constructed by Life  
 Technologies."  
 BASE COUNT 318 a 351 c 320 g 152 t  
 ORIGIN  
 Query Match 26.8%; Score 111.8; DB 10; Length 1141;  
 Best Local Similarity 94.3%; Pred. No. 2.5e-14;  
 Matches 116; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 Oy 1 ATGGGGGAACCTGCGCGAGGACTCCGCACTCCGCACTGAGAGAGAGACACTGTGG 60  
 Db 161 ATGGGGGAACCTGTGCGGAGGACTCCGCACTCCGCACTGAGAGAGAGACACTGTGG 220  
 Oy 61 GAGATGATGAGAGAGACCGGCAAGATCGTACGCTGATCGCCAGCGGCTCAAC 120  
 Db 221 GAGATGATGAGAGAGACCGGCAAGATCGTACGCTGATCGCCAGCGGCAATTAA 280  
 Oy 121 CCC 123  
 Db 281 CCC 283  
 RESULT 11  
 LOCUS CD349577 732 bp mRNA linear EST 29-MAY-2003  
 DEFINITION UT-M-FY0-cfr-k-23-0-UI-r1 NIH\_BMAP\_FY0 Mus musculus CDNA clone  
 IMAGE: 6855024 5', mRNA sequence.  
 ACCESSION CD349577  
 VERSION CD349577.1 GI:31141092  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 732)

AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: Dr. Jim Lin, University of Iowa  
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Distribution information can be found at  
<http://genome.uiowa.edu/distribution/mouse1.html>  
 This clone was contributed by the Brain Molecular Anatomy Project  
 (BMAP)  
 Seq primer: pYX-5.  
 Location/Qualifiers  
 1..732  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE: 6855024"  
 /tissue\_type="whole brain"  
 /dev\_stage="embryo 13.5,14.5,16.5,17.5dpc"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /clone\_1ib="NIH BMAP FY0"  
 /note="Organ: Brain; Vector: pYX-Asc; Site\_1: EcoR I;  
 Site\_2: Not I; The library was constructed according  
 Bonafide, Lennon and Soares, Genome Research, 6:791-806,  
 1996. Denatured RNA was size fractionated on a 1% agarose  
 gel. First strand cDNA synthesis was primed with oligo-dT  
 primer containing a Not I site. Double strand cDNA was  
 size selected according to mRNA size fraction, ligated  
 with EcoR I adaptor, digested with NotI and then cloned  
 directionally into pYX-Asc vector. The library tag  
 sequence located between the Not I site and the polyA tail  
 is AGCGAGACAG. This library was created for the University  
 Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the  
 Developing Mouse Nervous System', supported by National  
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,  
 program coordinator."  
 BASE COUNT 130 a 228 c 277 g 97 t  
 ORIGIN  
 Query Match 26.4%; Score 110.2; DB 14; Length 732;  
 Best Local Similarity 61.8%; Pred. No. 4.9e-14;  
 Matches 175; Conservative 0; Mismatches 108; Indels 0; Gaps 0;  
 Oy 43 GAGAGAGAGACACTGTGGGAGATGATGAGAGACCGGCAAGATCGTACGCTGATC 102  
 Db 309 GAGAGAGAGACCGTGTGGGAGCGGATGAGAGGGGTCCGACCGATGACCGCGCCCTG 368  
 Oy 103 TGCCCGAGCGGCTCAACCCCTTACTGCGCCAGGCCAAGTGTGTGCCAGCTGAGAG 162  
 Db 369 AATCCGCGCAAGTCAACCGCTTATCTGCGCCAGTCCGCGTCTTGAGACGAAGAGAG 428  
 Oy 163 GAGAGAGTGTGCAAGACCCCGGCTACCAACAGGCGCATGCGGGCCGAGCACTTGCTG 222  
 Db 429 GAGAGAGTGTGCAAGACCTTACGCTTCCGCTGCGGCTTAACTGACCGGCGCTCATTA 488  
 Oy 223 GATTGTGAAGACTGAGAGGAAGAACGGGGCCATGCTCTCTTGAGAGAGCTGAAATTC 282  
 Db 489 GACATTTTACGCTGCGGGGCAAGAGAGGGCTTTCGAGGCGCTTCTTGAAACCCCTGAAATTC 548  
 Oy 283 CACAAACCTGAGCTTCAACCCCTGATCAACCGGGCTGAGAGCTG 325  
 Db 549 TACTTACCAGAAACATTCAAGCTGTCTACTGAGCCAGAAACTG 591  
 RESULT 12  
 LOCUS BE287025 846 bp mRNA linear EST 26-OCT-2000  
 DEFINITION 601097271P1 NCI\_CGAP\_Mam5 Mus musculus CDNA clone IMAGE:3496121 5',  
 mRNA sequence.



ACCESSION BE287025  
 VERSION BE287025.1 GI:9165749  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE NIH-MGC <http://mgi.nci.nih.gov/>  
 1 (bases 1 to 846)  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Straubeberg, Ph.D.  
 Email: cga@bbs-remail.nih.gov  
 Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: LLAM8547 row: m column: 18  
 High quality sequence stop: 697.  
 Location/Qualifiers  
 1.846  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:3496121"  
 /tissue\_type="tumor, gross tissue"  
 /dev\_stage="7 months"  
 /lab\_host="DH10B"  
 /clone\_1lb="NCI CGAP Mam5"  
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1; Salt; Site: 2; Notti; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Robin Humphreys, NIH"  
 BASE COUNT 159 a 244 c 314 g 129 t  
 ORIGIN  
 Query Match 26.4%; Score 110.2; DB 10; Length 846;  
 Best Local Similarity 61.8%; Pred. No. 5.1e-14;  
 Matches 175; Conservative 0; Mismatches 108; Indels 0; Gaps 0;  
 QY 43 GACGAGAGACACTGTGGAGATGATGAGAGCCACGCCACAGATCGTACCTGCATC 102  
 DB 115 GAAGAGAGACGCGTGTGGAGCGATGAGAGCGGTCCGACGACTGACGGCGCCCTG 174  
 QY 103 TGCCCAAGCCGCTCACCCCTTACCTGCGCCAGGCCAAGGTGCTGTGCCAGTGAAG 162  
 DB 175 AATCCGGCCAAAGCTCAGCGCTTACTGCGCAGTGCCTCTGAGCAACAGAGAGAG 234  
 QY 163 GAGAGAGTGTGCAACACCCCGGCTCACCAAGCGCCATGGGCGCCGCGCACTTGCTG 222  
 DB 235 GAGGAGGTGTGCAAGACCTTACCTGTCCTGCGCGCTTACCGTACCGGCGCCCTCAT 294  
 QY 223 GATTGCTGAAGACTCGAGGAGAAAGCGGGCCATCGCTTCTGAGAGAGCTGAAGTTT 282  
 DB 295 GACATTTTACGCTCCCGGGGAGAGAGGGCTTGAAGGCTTCTTGAAGGCTTGAAATTC 354  
 QY 283 CACAACCTGACGTCTACACCTGTGTGACCGGGCTGACGCTG 325  
 DB 355 TACTACCAAGAACTTCACGCTGTCTACCTGGCCAGGAACCTG 397  
 RESULT 13  
 AJ455484 645 bp mRNA linear EST 22-APR-2002  
 LOCUS AJ455484  
 DEFINITION AJ455484 riken1 Gallus gallus cDNA clone 5k17r1, mRNA sequence.  
 ACCESSION AJ455484  
 VERSION AJ455484.1 GI:20265580  
 KEYWORDS EST.

SOURCE Gallus gallus (chicken)  
 ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Gallus.  
 REFERENCE Buerstedde, J.M.  
 1 (bases 1 to 645)  
 TITLE Gallus gallus bursal lymphocyte EST  
 JOURNAL Unpublished  
 COMMENT Contact: Buerstedde JM  
 Cellular Immunology  
 Heinrich-Pette-Institute  
 Martinstr. 52, 20251 Hamburg, Germany  
 Email: URL: <http://genetics.hpi.uni-hamburg.de/dt40est.html>  
 Location/Qualifiers  
 1.645  
 /organism="Gallus gallus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9031"  
 /clone="5k17r1"  
 /cell\_type="bursal lymphocyte"  
 /dev\_stage="2-3 weeks old"  
 /clone\_1lb="riken1"  
 /note="CB inbred strain"  
 BASE COUNT 166 a 164 c 172 g 142 t 1 others  
 ORIGIN  
 Query Match 25.6%; Score 106.6; DB 9; Length 645;  
 Best Local Similarity 61.2%; Pred. No. 2.9e-13;  
 Matches 172; Conservative 0; Mismatches 109; Indels 0; Gaps 0;  
 QY 43 GACGAGAGACACTGTGGAGATGATGAGAGCCACGCCACAGATCGTACCTGCATC 102  
 DB 224 GAGGAGAGACTTGTGTGGAGAAAGTCAAGTGCACACCGGCACATGCTGACGCTTACATC 353  
 QY 103 TGCCCAAGCCGCTCACCCCTTACCTGCGCCAGGCCAAGGTGCTGTGCCAGTGAAG 162  
 DB 354 AACCAAGTAAAGCTGACCCCGTACCTGCGGCAATGCAAGTATGATGACCAAGATGAA 413  
 QY 163 GAGAGGTGTGTCACAGCCCGGCTCACCAAGCCGATGCGGCGGCGGCACTTGCTG 222  
 DB 414 GATGAGGTGTCTTACTCACTTATGCTGCTCCCAAAATTAACGAGCGCGGCTGCTG 473  
 QY 223 GATTGCTGAAGACTCGAGGAGAAAGCGGGCCATCGCTTCTGAGAGGCTGAAGTTT 282  
 DB 474 GACATTTCTTACACCAAGGCGAGAGGGCTATGTGTTTCTTGAAGAGCTTGAGATT 533  
 QY 283 CACAACCTGACGTCTTACACCTGTGTGTCACCGGCTGACGCC 323  
 DB 534 TACTACCTGAACCTTACAACTGTGTGCAAGAGAGAGGCC 574  
 RESULT 14  
 AJ453294 769 bp mRNA linear EST 22-APR-2002  
 LOCUS AJ453294  
 DEFINITION AJ453294 riken1 Gallus gallus cDNA clone 33024r1, mRNA sequence.  
 ACCESSION AJ453294  
 VERSION AJ453294.1 GI:20263390  
 KEYWORDS EST.  
 SOURCE Gallus gallus (chicken)  
 ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Gallus.  
 REFERENCE Buerstedde, J.M.  
 1 (bases 1 to 769)  
 TITLE Gallus gallus bursal lymphocyte EST  
 JOURNAL Unpublished  
 COMMENT Contact: Buerstedde JM  
 Cellular Immunology  
 Heinrich-Pette-Institute  
 Martinstr. 52, 20251 Hamburg, Germany  
 Email: URL: <http://genetics.hpi.uni-hamburg.de/dt40est.html>

FEATURES  
source

Location/Qualifiers  
1..769  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9031"  
/clone="33024r1"  
/cell\_type="bursal lymphocyte"  
/dev\_stage="2-3 weeks old"  
/clone\_1lb="rikeni"  
/note="CB inbred strain"

BASE COUNT 200 a 193 c 198 g 176 t 2 others

ORIGIN

Query Match 25.6%; Score 106.6; DB 9; Length 769;  
Best Local Similarity 61.2%; Pred. No. 3e-13;  
Matches 172; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 43 GAGAGAGACACGCTGGAGATGATGAGAGCCACAGGATCGTACGTCATC 102  
DB 315 GAGAGAGAGCTTTGTGGAGATGTCAGTGCACCGGCACATGTCAGCCCTTACATC 374  
QY 103 TGCCCGACCGCCTCACCCCTTACCTGGCCAGGCTGTGTCAGTGCAGAG 162  
DB 375 AACCCAGCTAGCTGAGCCCGCTACCTGGGCAATGCATGACCAAGATGAA 434  
QY 163 GAGAGGTGTCACAGCCCGGCTCACCAAGCCGATCGCGGCGGCACTTGTG 222  
DB 435 GATGAGGTGCTTAATCTCATCTATGCTGCGCTCCAAATTAACGAGGCGGCTGCTG 494  
QY 223 GATTGCTGAAGACTCGAGGAGAAAGCGGGCCATCGCCTTCTGAGAGCTTGAATTC 282  
DB 495 GACATTTCTTACACCAAGGCGCAGAGGCGCTATGTGTTCTTGAGAGCTTGAATTT 554  
QY 283 CACAACCTGAGCTTACACCCCTGATCAACCGGCTGAGGCC 323  
DB 555 TACTACCTGAACTTACAACTGTGTGACAGGGAAGGCC 595

RESULT 15  
AJ455489 584 bp mRNA linear EST 22-APR-2002

LOCUS AJ455489 riken1 Gallus gallus cDNA clone 5k23r1, mRNA sequence.  
DEFINITION AJ455489  
ACCESSION AJ455489  
VERSION AJ455489.1 GI:20265585  
KEYWORDS EST.

SOURCE  
ORGANISM Gallus gallus (chicken)  
Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
1 (bases 1 to 584)  
Buerstedde, J.W.  
Buerstedde, J.W.  
Gallus gallus bursal lymphocyte EST  
Unpublished  
Contact: Buerstedde JM  
Cellular Immunology  
Heinrich-Pette-Institute  
Martinistr. 52, 20251 Hamburg, Germany  
Email: URI: <http://genetics.hpi.uni-hamburg.de/dt40est.html>.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Location/Qualifiers  
1..584  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9031"  
/clone="5k23r1"  
/cell\_type="bursal lymphocyte"  
/dev\_stage="2-3 weeks old"  
/clone\_1lb="rikeni"  
/note="CB inbred strain"

FEATURES  
source

BASE COUNT 151 a 152 c 155 g 122 t 4 others

ORIGIN

Query Match 25.4%; Score 105.8; DB 9; Length 584;

Best Local Similarity 60.5%; Pred. No. 4.1e-13;  
Matches 170; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 43 GAGAGAGACACGCTGGAGATGATGAGAGCCACAGGATCGTACGTCATC 102  
DB 294 GAGAGAGAGCTTTGTGGAGATGTCAGTGCACCGGCACATGTCAGCCCTTACATC 353  
QY 103 TGCCCGACCGCCTCACCCCTTACCTGGCCAGGCTGTGTCAGTGCAGAG 162  
DB 354 AACCCAGCTAGCTGAGCCCGCTACCTGGGCAATGCATGACCAAGATGAA 413  
QY 163 GAGAGGTGTCACAGCCCGGCTCACCAAGCCGATCGCGGCGGCACTTGTG 222  
DB 414 GATGAGGTGCTTAATCTCATCTATGCTGCGCTCCAAATTAACGAGGCGGCTGCTG 473  
QY 223 GATTGCTGAAGACTCGAGGAGAAAGCGGGCCATCGCCTTCTGAGAGCTTGAATTC 282  
DB 474 GACATTTCTTACACCAAGGCGCAGAGGCGCTATGTGTTCTTGAGAGCTTGAATTT 533  
QY 283 CACAACCTGAGCTTACACCCCTGATCAACCGGCTGAGGCC 323  
DB 534 TACTACCTGAACTTACAACTGTGTGACAGGGAAGGCC 574

Search completed: February 17, 2004, 23:57:11  
Job time : 2349.94 secs



Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 12709  
LENGTH: 2430  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (91)  
OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.  
US-09-252-991A-12709

Query Match 10.5%; Score 43.6; DB 4; Length 2430;

Best Local Similarity 48.8%; Pred. No. 0.12;

Matches 118; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 80 GCCACGAGATCGTACGCTGATCTGCGCCGCTCACTCCCTACCTGCGCGACGCA 139  
DB 1193 GCAAGACGACGATCGGCAATCATCGCCGAGAGCTCGCGCGCTTCACTGTTCA 1252  
QY 140 AGGTGCTGTGCTGACGCTGACGAGAGAGAGTGTGACAGCCCGGCTCAGCAACAGCG 199  
DB 1253 GCGTCGCGCGGCGACGCGCGAGGAGATCAAGGCGCACCGCGCATCTACGCGCG 1312  
QY 200 CCATGCGGCGCGGCGGCGCTGCTGATTTGCTGAAGCTCGAGGGAAGACGGGCGCATCG 259  
DB 1313 CCTGCGCGCGGCGGCGGCTGCTGAGGAGGAGTGAAGAGTGAAGTGAAGTGAAGTGA 1372  
QY 260 CCTTCTGAGAGGCTGAGATTTCACACCTGACGCTTACACCTGCTCACCGGAGTGC 319  
DB 1373 TGCTGACGAGATCGACAGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCTGC 1432  
QY 320 AG 321  
DB 1433 TG 1434

RESULT 3  
US-09-252-991A-13249/c  
Sequence 13249, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 13249  
LENGTH: 2607  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (2532)  
OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.  
US-09-252-991A-13249

Query Match 10.5%; Score 43.6; DB 4; Length 2607;  
Best Local Similarity 48.8%; Pred. No. 0.12;  
Matches 118; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 80 GCCACGAGATCGTACGCTGATCTGCGCCGCTCACTCCCTACCTGCGCGACGCA 139  
DB 1430 GCAAGACGACGATCGGCAATCATCGCCGAGAGCTCGCGCGCTTCACTGTTCA 1371  
QY 140 AGGTGCTGTGCTGACGCTGACGAGAGAGTGTGACAGCCCGGCTCAGCAACAGCG 199  
DB 1373 GCGTCGCGCGGCGGCGGCGGCTGCTGAAGCTCGAGGGAAGACGGGCGCATCG 1311  
QY 200 CCATGCGGCGCGGCGGCGGCTGCTGATTTGCTGAAGCTCGAGGGAAGACGGGCGCATCG 259  
DB 1310 CCTGCGCGCGGCGGCGGCTGCTGAGGAGTGAAGAGTGAAGTGAAGTGAAGTGAAGTGA 1251  
QY 260 CCTTCTGAGAGGCTGAGATTTCACACCTGACGCTTACACCTGCTCACCGGAGTGC 319  
DB 1250 TGCTGACGAGATCGACAGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCTGC 1191  
QY 320 AG 321  
DB 1190 TG 1189

RESULT 4  
US-08-533-669A-7/c

Sequence 7, Application US/08533669A

Patent No. 5834592

GENERAL INFORMATION:

APPLICANT: Corixa Corporation

TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE

TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/533,669A

FILING DATE: 22-SEP-1995

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Mark, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 210121,420

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 1771 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1698

US-08-533-669A-7

Query Match 10.4%; Score 43.4; DB 2; Length 1771;

Best Local Similarity 46.5%; Pred. No. 0.13; Indels 0; Gaps 0;

Matches 140; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

QY 19 AGGAGTCCGGCACTACCGGCACTGAGCCAGAGAGACACTGTGGAGATGATGAGACAC 78  
Db 432 ACGCGCTGCTCTCTTCAGAGGCGCTTGACGCGGGCGCTGCTGGCGCGATGATGATCATTCAG 373  
QY 79 CGCCACAGAGANTGATAGCTGTCATCTGAGCCCAAGCGGCTCACCCCTCACTGCGCAGAGCC 138  
Db 372 GCGGGGCGGCTCTCTGGCTTCTCTCAAGCTCCGCGCGGAGACGGCTGCTCTTGAGGGC 313  
QY 139 AAGGTCTGTGCCAGCTGAGTCGACGAGAGAGAGTGTGCAAGACCCCGGCTTCACCAACGC 198  
Db 312 CTGGAGAGCGGGCGCTGCTCGGCGCTCGAGATGAGCTTCAGAGCGCGCGGCTCTCTCGGACGC 253  
QY 199 GCCATGCGGGCGCGGCACTTGTCTGATTTGTGAAGACTCGAGGGAGAGACGGGGCCATC 258  
Db 252 CTCACAGCTCCGCGCGAGAGACGCGCTGCTCTCGAGGGCGCTGACGCGAGGCTCTCGGC 193  
QY 259 GCGTTTCTGGAGAGGCTGAAATTCCAAACCTGACGCTTACACCTGTGTAACCGGGCTG 318  
Db 192 CGCATGATGACATTCAGGCGGGCGCGCTCTCGCGCTCTTCACAGCTCCGCGCGAGACG 133  
QY 319 C 319  
Db 132 C 132

## RESULT 5

```

US-08-511-872-1/c
Sequence 1, Application US/08511872
Patent No. 5965142
GENERAL INFORMATION:
APPLICANT: Dillon, Davin C.
APPLICANT: Reed, Steven G.
APPLICANT: Day, Craig H.
TITLE OF INVENTION: POLYPEPTIDES AND METHODS FOR THE
TITLE OF INVENTION: DETECTION OF L. tropica INFECTION
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY
STREET: 6300 Columbia Avenue, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 94104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/511,872
FILING DATE: 04-AUG-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: NO. 5965142tenbury, Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 210121.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
TELEX: 3723836 SEEDNDBERRY
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1771 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1698
US-08-511-872-1

```

Query Match	10.4%;	Score 43.4;	DB 2;	Length 1771;
Best Local Similarity	46.5%;	Pred. No. 0.13;		

	Matches	140;	Conservative	0;	Mismatches	161;	Indels	0;	Gaps	0;
QY	19	AGGACCTTCGACCTCAGCGCACCTGACGAGAGACACTGTGGAGATGAGAGCCAC	78							
Db	432	ACGGGCTCCCTCCCTCGAGGCGCTTGAGACGCGGGCGCTGCTCGGCGCGCATGACATCCAG	373							
QY	79	CGCCACAGAGATCGTACGCTGATCTGGCCGACCGGCGCTCACCCCTACTGTGGCCAGGCG	138							
Db	372	GCGGCGCGCCCTCCTCGGCTTCTCCAGTCTCGCGCGGAGAGCGGTGTCTCTCCACAGGCG	313							
QY	139	AAGGTGCTGTGCTCCGACCTGACGAGAGAGAGGTGCTGACAGCCCGGCTCAACCAAGC	198							
Db	312	CTGGGAGGGGCGCTGCTCGGCTCGTGTGATAGGCTTCCAGGGCGCGGCTCTCTGGGACGC	253							
QY	199	GCCATGCGGGCGCGGACCTTGCTGGAACTCGAGGAGAAACCGGGCCATC	258							
Db	252	CTCCAGCTCCGCGGGAGAGCGCGGTGCTCTCCAGGGCGCTGAGACGGGGCTGTGCTCGGC	193							
QY	259	GCCTTCTCTGAGAGCGCTGAAATTCCACACCCCTGAAGTCAACCTTGGTCAACGGGCTG	318							
Db	192	CGCATGTGACATCCAGCGCGGCGCGCTCTCTCGGCTCTCTCCAGCTCCGCGCGGAGAGC	133							
QY	319	C 319								
Db	132	C 132								

## RESULT 6

US-09-183-861-7/C  
Sequence 7, Application US/09183861  
Patent No. 6365165

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.  
APPLICANT: Campos-Neto, Antonio  
APPLICANT: Webb, John R.  
APPLICANT: Dillon, David C.  
APPLICANT: Skelky, Valet A.W.  
TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE THERAPY AND  
NUMBER OF SEQUENCES: 87

CORRESPONDENCE ADDRESSES:

ADDRESSES: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/183,861  
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/022,765  
FILING DATE: 12-FEB-1998

ATTORNEY/AGENT INFORMATION:

NAME: MAKI, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.420C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 1771 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1698  
US-09-183-861-7

Query Match 10.4%; Score 43.4; DB 4; Length 1771;  
Best Local Similarity 46.5%; Pred. No. 0.13;  
Matches 140; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

QY 19 AGGAGCTCCGCACTCAGGCACTGAGAGAGAGACACTGTGGAGATGATGAGAGCCAC 78  
DB 432 AGCGCTGCTCTCTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 373  
QY 79 CGGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 138  
DB 372 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 313  
QY 139 AAGGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 198  
DB 312 CTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 253  
QY 199 GCCATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 258  
DB 252 CTCAGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 193  
QY 259 GCTTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 318  
DB 192 CGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 133  
QY 319 C 319  
DB 132 C 132

RESULT 7  
US-09-022-765-7/c

Sequence 7, Application US/09022765  
Patent No. 6375955

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.  
APPLICANT: Campos-Neto, Antonio  
APPLICANT: Webb, John R.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Skeiky, Yasir A.M.  
TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE THERAPY AND  
NUMBER OF SEQUENCES: 87  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
City: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/022,765  
FILING DATE: 12-FEB-1998  
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Marki, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 210121.420C3

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 1771 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1698  
US-09-022-765-7

Query Match 10.4%; Score 43.4; DB 4; Length 1771;  
Best Local Similarity 46.5%; Pred. No. 0.13;  
Matches 140; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

QY 19 AGGAGCTCCGCACTCAGGCACTGAGAGAGAGACACTGTGGAGATGATGAGAGCCAC 78  
DB 432 AGCGCTGCTCTCTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 373  
QY 79 CGGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 138  
DB 372 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 313  
QY 139 AAGGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 198  
DB 312 CTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 253  
QY 199 GCCATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 258  
DB 252 CTCAGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 193  
QY 259 GCTTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 318  
DB 192 CGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 133  
QY 319 C 319  
DB 132 C 132

RESULT 8

US-09-551-974A-7/c

Sequence 7, Application US/09551974A

Patent No. 650437

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.  
APPLICANT: Campos-Neto, Antonio  
APPLICANT: Webb, John R.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Skeiky, Yasir A.M.  
TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE  
TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS  
FILE REFERENCE: 210121.420C5  
CURRENT APPLICATION NUMBER: US/09/551,974A  
CURRENT FILING DATE: 2000-04-14  
NUMBER OF SEQ ID NOS: 101  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 7  
LENGTH: 1771  
TYPE: DNA  
ORGANISM: Leishmania tropica  
US-09-551-974A-7

Query Match 10.4%; Score 43.4; DB 4; Length 1771;  
Best Local Similarity 46.5%; Pred. No. 0.13;  
Matches 140; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

QY 19 AGGAGCTCCGCACTCAGGCACTGAGAGAGAGACACTGTGGAGATGATGAGAGCCAC 78  
DB 432 AGCGCTGCTCTCTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 373  
QY 79 CGGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 138  
DB 372 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 313  
QY 139 AAGGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 198  
DB 312 CTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 253





GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 351  
LENGTH: 2016  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-351

Query Match 10.3%; Score 42.8; DB 4; Length 2016;  
Best Local Similarity 51.6%; Pred. No. 0.18;  
Matches 98; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 59 GGGAGATGATGAGAGACCAACCGGACAGATGTAAGTCACTGACCCAGCGGCTCA 118  
DB 623 GGAAGCGGCTGTGACCTGATGCTCCGAAATGCTCGGCGGCTTCAAGCGCGCATCG 564  
QY 119 CCCCCTACCTGCGCCAGCGCAAGTGTGCTGCGAGCGAGAGAGAGTGTGACA 178  
DB 563 ATACCAACGTCGCGCGCGGCGCAAGCGACGCTGCGCGGCTTGAAGAGTGTGCGGTA 504  
QY 179 GCGCCCGGCTCACCAACAGCGGCTGCGGCGGCGGCACTTGTGATTGTGAAGACTC 238  
DB 503 TCGCCACGTTGCGCCCGGCTGCGAGGTGCGAGCGCGGCTGCAACAGCTC 444  
QY 239 GAGGGAAGAA 248  
DB 443 GTTCGAAGGA 434

## RESULT 13

US-09-252-991A-5873  
Sequence 5873, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 5873  
LENGTH: 534  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-5873

Query Match 10.2%; Score 42.6; DB 4; Length 534;  
Best Local Similarity 54.9%; Pred. No. 0.15;  
Matches 84; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 99 CATCTGCCCCAGCGGCTCAACCCCTACCTGCGCGCAAGTGTGCTGCAAGTGA 158  
DB 337 CAGCGCGGACCTGCTTCTCCGCGCCTTGCAGAGCGGAGCTGCTGCGGCAAGCC 396  
QY 159 CGAGAGAGAGTGTCTCAACGCGGCTCAACCAAGCGGCTGCGGCGGCGGCACTT 218  
DB 397 GCTGAGAGAGGCGCTGAAAGGCGTGTCTTCAATGAGATGAGAGCGAGCAAGACTCT 456

QY 219 GCTGATTGTGTAAGACTGAGGAGAAAGG 251  
DB 457 GCTAGCGCGCGCGGCGCGCGCGCAAGGCGG 489

## RESULT 14

US-09-252-991A-5829  
Sequence 5829, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 5829  
LENGTH: 1968  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-5829

Query Match 10.2%; Score 42.6; DB 4; Length 1968;  
Best Local Similarity 54.9%; Pred. No. 0.2;  
Matches 84; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 99 CATCTGCCCCAGCGGCTCAACCCCTACCTGCGCGCAAGTGTGCTGCAAGTGA 158  
DB 1812 CAGCGCGGACCTGCTTCTCCGCGCCTGCGCAAGCGGAGTGTGCGCGGCAAGCC 1871  
QY 159 CGAGAGAGAGTGTCTCAACGCGGCTGACCAAGCGGCTGCGGCGGCGGCACTT 218  
DB 1872 GCTGAGAGAGCGCTGAAAGCGGCTGCTTCAATGAGATGAGAGCGGAGCAAGCTCT 1931  
QY 219 GCTGATTGTGTAAGACTGAGGAGAAAGG 251  
DB 1932 GCTAGCGCGCGCGGCGCGCGCGCAAGGCGG 1964

## RESULT 15

US-09-252-991A-5779/c  
Sequence 5779, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 5779  
LENGTH: 2190  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-5779

Query Match 10.2%; Score 42.6; DB 4; Length 2190;  
Best Local Similarity 54.9%; Pred. No. 0.2;  
Matches 84; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 99 CATCTGCCCCAGCGGCTCAACCCCTACCTGCGCGCAAGTGTGCTGCAAGTGA 158  
DB 160 CAGCGCGGACCTGCTTCTCCGCGCCTGCGCAAGCGGAGCTGCTGCGGCAAGCC 101

QY 159 CGAGAGAGAGTGTGTGACAGCCCGGCTCACCAACAGCGCCATGCGGCGCGGCACTT 218  
 Db 100 GCTGAGAGAGGCGCTTGAAAGCGGTCCCTTCATGAGATGACGACCGGACGACTCTCT 41  
 QY 219 GCTGATTGCTGAGACTCGAGGGAAGAACGG 251  
 Db 40 GCTGAGCGGCGCGGCGCGCGGCAAGGCGG 8

Search completed: February 17, 2004, 23:58:57  
 Job time : 61.1645 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 17, 2004, 22:52:28 (Search time 269.576 Seconds  
(without alignments)  
5698.119 Million cell updates/sec

Title: US-10-032-159A-15

Sequence: 1 atgggggaactgtgcgcag.....ggctctcgtaaccacagm 417

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapept 1.0

Searched: 244703 segs, 1841816367 residues

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Published Applications NA.\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	415	99.5	417	14	US-10-032-159A-15
2	415	99.5	3417	9	US-09-767-215-4
3	351	84.2	3012	9	US-09-767-215-3
4	351	84.2	3766	12	US-10-104-047-315
5	351	84.2	3951	9	US-09-767-215-1
6	276	66.2	276	14	US-10-032-159A-17
7	117.8	28.2	3096	11	US-09-798-412-9
8	117.8	28.2	3096	15	US-10-325-917-9
9	117.8	28.2	3949	11	US-09-798-412-7
10	117.8	28.2	3949	15	US-10-325-917-7
11	111.8	26.8	1141	14	US-10-032-159A-37
12	104	24.9	3441	11	US-09-798-412-12
13	104	24.9	3441	15	US-10-325-917-12
14	104	24.9	3744	14	US-10-032-159A-7
15	104	24.9	4276	11	US-09-798-412-10

16	104	24.9	4276	15	US-10-325-917-10	Sequence 10, Appl
17	100.4	24.1	276	14	US-10-032-159A-9	Sequence 9, Appl
18	94.4	22.6	1101	14	US-10-032-159A-1	Sequence 1, Appl
19	94.4	22.6	1608	11	US-09-798-412-6	Sequence 6, Appl
20	94.4	22.6	1608	15	US-10-325-917-6	Sequence 6, Appl
21	94.4	22.6	2098	11	US-09-798-412-4	Sequence 4, Appl
22	94.4	22.6	2098	15	US-10-325-917-4	Sequence 4, Appl
23	94.4	22.6	2176	14	US-10-032-159A-19	Sequence 19, Appl
24	94	22.5	514	10	US-09-796-692-3560	Sequence 3560, Ap
25	94	22.5	514	12	US-10-057-4758-3560	Sequence 3560, Ap
26	94	22.5	514	12	US-10-154-8848-3560	Sequence 3560, Ap
27	94	22.5	514	15	US-10-040-862-3560	Sequence 3560, Ap
28	85.2	20.4	1608	11	US-09-798-412-3	Sequence 3, Appl
29	85.2	20.4	1608	15	US-10-325-917-3	Sequence 3, Appl
30	85.2	20.4	1879	11	US-09-798-412-1	Sequence 1, Appl
31	85.2	20.4	1879	15	US-10-325-917-1	Sequence 1, Appl
32	69.8	16.7	216	14	US-10-032-159A-3	Sequence 3, Appl
33	51.8	12.4	281	14	US-10-032-159A-21	Sequence 21, Appl
34	51.8	12.4	500	13	US-10-029-386-12805	Sequence 12805, A
35	49.8	11.9	804	13	US-10-027-632-172688	Sequence 172688,
36	49.8	11.9	804	14	US-10-027-632-172688	Sequence 172688,
37	49	11.8	858	13	US-10-027-632-166977	Sequence 166977,
38	49	11.8	858	14	US-10-027-632-166977	Sequence 166977,
39	48	11.5	145	13	US-10-029-386-26505	Sequence 26505, A
40	47	11.3	509	13	US-10-027-632-290760	Sequence 290760,
41	47	11.3	509	13	US-10-027-632-290761	Sequence 290761,
42	47	11.3	509	14	US-10-027-632-290760	Sequence 290761,
43	47	11.3	509	14	US-10-027-632-290761	Sequence 290761,
44	45	10.8	36620	11	US-09-952-060-30	Sequence 30, Appl
45	44.6	10.7	1548	12	US-10-369-493-31289	Sequence 31289, A

ALIGNMENTS

RESULT 1  
US-10-032-159A-15  
Sequence 15, Application US/10032159A  
Publication No. US20020164703A1  
GENERAL INFORMATION:  
APPLICANT: Pawlowski, Krzysztof  
APPLICANT: Reed, John C.  
TITLE OF INVENTION: CARD-DOMAIN CONTAINING POLYPEPTIDES.  
TITLE OF INVENTION: ENCODING NUCLEIC ACIDS, AND METHODS OF USE  
FILE REFERENCE: P-LJ 5100  
CURRENT APPLICATION NUMBER: US/10/032,159A  
CURRENT FILING DATE: 2001-12-19  
PRIOR APPLICATION NUMBER: US 60/257,457  
PRIOR FILING DATE: 2000-12-21  
NUMBER OF SEQ ID NOS: 37  
SOFTWARE: FastSeq For Windows Version 4.0  
SEQ ID NO 15  
LENGTH: 417  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1) ... (417)  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 416..417  
OTHER INFORMATION: n = A,T,C or G  
US-10-032-159A-15  
Query Match 99.5% Score 415; DB 14; Length 417;  
Best Local Similarity 100.0%; Pred. No. 1.1e-107; Indels 0; Gaps 0;  
Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGGGAACTGTGCGGAGGACTCGGCACTGACGAGCACTGAGAGAGACTGTGG 60  
DB 1 ATGGGGAACTGTGCGGAGGACTCGGCACTGACGAGCACTGAGAGAGACTGTGG 60

QY 61 GAGATGATGAGAGAGCCACCGCCACAGATGCTACCGCTGATCTGCCACGCGCTCAC 120  
Db 61 GAGATGATGAGAGAGCCACCGCCACAGATGCTACCGCTGATCTGCCACGCGCTCAC 120  
QY 121 CCTACTGCGCGCAGGCGCAAGGTGCTGTGTCAGCTGAGCAGAGAGAGAGTGTGCAAGC 180  
Db 121 CCTACTGCGCGCAGGCGCAAGGTGCTGTGTCAGCTGAGCAGAGAGAGAGTGTGCAAGC 180  
QY 181 CCGCGGCTCACCAACAGCGCCATGCGGCGCGGCACTTGCTGATTTGCTGAAGACTGA 240  
Db 181 CCGCGGCTCACCAACAGCGCCATGCGGCGCGGCACTTGCTGATTTGCTGAAGACTGA 240  
QY 241 GGGAGAGAGCGGGCCATCGCTTCTCTGAGAGCTGGAAGTTCCAAACCTGACCTTAC 300  
Db 241 GGGAGAGAGCGGGCCATCGCTTCTCTGAGAGCTGGAAGTTCCAAACCTGACCTTAC 300  
QY 301 ACCCTGATGACCGGCGCTGAGGCTGATGTTGACTTCAAGTAAGTTAGCGTGAGAGCTCC 360  
Db 301 ACCCTGATGACCGGCGCTGAGGCTGATGTTGACTTCAAGTAAGTTAGCGTGAGAGCTCC 360  
QY 361 GACTTGACGCTTTGGCAGGCACTTCTAGGAAGCTCAGGCTCTGTGTAACCCAG 415  
Db 361 GACTTGACGCTTTGGCAGGCACTTCTAGGAAGCTCAGGCTCTGTGTAACCCAG 415

## RESULT 2

US-09-767-215-4

; Sequence 4, Application US/09767215

; Patent No. US20020081636A1

; GENERAL INFORMATION:

; APPLICANT: Berlin, John

; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED

; FILE REFERENCE: 07334-142001

; CURRENT APPLICATION NUMBER: US/09/767,215

; CURRENT FILING DATE: 2001-01-22

; PRIOR APPLICATION NUMBER: 60/181,159

; PRIOR FILING DATE: 2000-02-09

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 3417

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)...(3417)

; US-09-767-215-4

Query Match 99.5%; Score 415; DB 9; Length 3417;

Best Local Similarity 100.0%; Pred. No. 1.5e-107; Indels 0; Gaps 0;

Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGGAACTGTGCGCGCAGGAGCTCCGACTCAGCGGCACTGAGAGAGAGCACTGTGG 60  
Db 1 ATGGGGGAACTGTGCGCGCAGGAGCTCCGACTCAGCGGCACTGAGAGAGAGCACTGTGG 60  
QY 61 GAGATGATGAGAGAGCCACCGCCACAGATGCTACGCTGATCTGCCACGCGGCTCAC 120  
Db 61 GAGATGATGAGAGAGCCACCGCCACAGATGCTACGCTGATCTGCCACGCGGCTCAC 120  
QY 121 CCTACTGCGCGCAGGCGCAAGGTGCTGTGTCAGCTGAGCAGAGAGAGAGTGTGCAAGC 180  
Db 121 CCTACTGCGCGCAGGCGCAAGGTGCTGTGTCAGCTGAGCAGAGAGAGAGTGTGCAAGC 180  
QY 181 CCGCGGCTCACCAACAGCGCCATGCGGCGCGGCACTTGCTGATTTGCTGAAGACTGA 240  
Db 181 CCGCGGCTCACCAACAGCGCCATGCGGCGCGGCACTTGCTGATTTGCTGAAGACTGA 240  
QY 241 GGGAGAGAGCGGGCCATCGCTTCTCTGAGAGCTGGAAGTTCCAAACCTGACCTTAC 300  
Db 241 GGGAGAGAGCGGGCCATCGCTTCTCTGAGAGCTGGAAGTTCCAAACCTGACCTTAC 300

QY 301 ACCCTGATGACCGGCGCTGAGGCTGATGTTGACTTCAAGTAAGTTAGCGTGAGAGCTCC 360  
Db 301 ACCCTGATGACCGGCGCTGAGGCTGATGTTGACTTCAAGTAAGTTAGCGTGAGAGCTCC 360  
QY 361 GACTTGACGCTTTGGCAGGCACTTCTAGGAAGCTCAGGCTCTGTGTAACCCAG 415  
Db 361 GACTTGACGCTTTGGCAGGCACTTCTAGGAAGCTCAGGCTCTGTGTAACCCAG 415

## RESULT 3

US-09-767-215-3

; Sequence 3, Application US/09767215

; Patent No. US20020081636A1

; GENERAL INFORMATION:

; APPLICANT: Berlin, John

; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED

; FILE REFERENCE: 07334-142001

; CURRENT APPLICATION NUMBER: US/09/767,215

; CURRENT FILING DATE: 2001-01-22

; PRIOR APPLICATION NUMBER: 60/181,159

; PRIOR FILING DATE: 2000-02-09

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 3012

; TYPE: DNA

; ORGANISM: Homo sapiens

; US-09-767-215-3

Query Match 84.2%; Score 351; DB 9; Length 3012;

Best Local Similarity 100.0%; Pred. No. 1.8e-89; Indels 0; Gaps 0;

Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGGAACTGTGCGCGCAGGAGCTCCGACTCAGCGGCACTGAGAGAGAGCACTGTGG 60  
Db 1 ATGGGGGAACTGTGCGCGCAGGAGCTCCGACTCAGCGGCACTGAGAGAGAGCACTGTGG 60  
QY 61 GAGATGATGAGAGAGCCACCGCCACAGATGCTACGCTGATCTGCCACGCGGCTCAC 120  
Db 61 GAGATGATGAGAGAGCCACCGCCACAGATGCTACGCTGATCTGCCACGCGGCTCAC 120  
QY 121 CCTACTGCGCGCAGGCGCAAGGTGCTGTGTCAGCTGAGCAGAGAGAGAGTGTGCAAGC 180  
Db 121 CCTACTGCGCGCAGGCGCAAGGTGCTGTGTCAGCTGAGCAGAGAGAGAGTGTGCAAGC 180  
QY 181 CCGCGGCTCACCAACAGCGCCATGCGGCGCGGCACTTGCTGATTTGCTGAAGACTGA 240  
Db 181 CCGCGGCTCACCAACAGCGCCATGCGGCGCGGCACTTGCTGATTTGCTGAAGACTGA 240  
QY 241 GGGAGAGAGCGGGCCATCGCTTCTCTGAGAGCTGGAAGTTCCAAACCTGACCTTAC 300  
Db 241 GGGAGAGAGCGGGCCATCGCTTCTCTGAGAGCTGGAAGTTCCAAACCTGACCTTAC 300  
QY 301 ACCCTGATGACCGGCGCTGAGGCTGATGTTGACTTCAAGTAAGTTAGCGTGAGAGCTCC 360  
Db 301 ACCCTGATGACCGGCGCTGAGGCTGATGTTGACTTCAAGTAAGTTAGCGTGAGAGCTCC 360

## RESULT 4

US-10-104-047-315

; Sequence 315, Application US/10104047

; Publication No. US20030236392A1

; GENERAL INFORMATION:

; APPLICANT: HELIX RESEARCH INSTITUTE

; TITLE OF INVENTION: NO. US20030236392A1 full length cDNA

; FILE REFERENCE: HI-A0105

; CURRENT APPLICATION NUMBER: US/10/104,047

; PRIOR APPLICATION NUMBER: 2002-03-25

; PRIOR FILING DATE:

; NUMBER OF SEQ ID NOS: 4096

; SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 315  
LENGTH: 3766  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-104-047-315

Query Match 84.2%; Score 351; DB 12; Length 3766;  
Best Local Similarity 100.0%; Pred. No. 1.8e-89;  
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGGAACTGTGCGCCAGGGACTCCGCACTCAGCGCACTGAGAGAGAGACTGTGG 60  
DB 192 ATGGGGGAACTGTGCGCCAGGGACTCCGCACTCAGCGCACTGAGAGAGAGACTGTGG 251  
QY 61 GAGATGATGAG 120  
DB 252 GAGATGATGAG 311  
QY 121 CCTTACCTGCGCCAGGAG 180  
DB 312 CCTTACCTGCGCCAGGAG 371  
QY 181 CCCCAGGCTCAGCAAG 240  
DB 372 CCCCAGGCTCAGCAAG 431  
QY 241 GGGAG 300  
DB 432 GGGAG 491  
QY 301 ACCCTGTGTCACCGGAG 351  
DB 492 ACCCTGTGTCACCGGAG 542

## RESULT 5

US-09-767-215-1  
Sequence 1, Application US/09767215  
Patent No. US20020081636A1  
GENERAL INFORMATION:  
APPLICANT: Bertin, John  
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED  
TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF  
FILE REFERENCE: 07334-142001  
CURRENT APPLICATION NUMBER: US/09/767,215  
CURRENT FILING DATE: 2001-01-22  
PRIOR APPLICATION NUMBER: 60/181,159  
PRIOR FILING DATE: 2000-02-09  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 3931  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (207)...(3218)  
US-09-767-215-1

Query Match 84.2%; Score 351; DB 9; Length 3911;  
Best Local Similarity 100.0%; Pred. No. 1.9e-89;  
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGGAACTGTGCGCCAGGGACTCCGCACTCAGCGCACTGAGAGAGAGACTGTGG 60  
DB 207 ATGGGGGAACTGTGCGCCAGGGACTCCGCACTCAGCGCACTGAGAGAGAGACTGTGG 266  
QY 61 GAGATGATGAG 120  
DB 267 GAGATGATGAG 326  
QY 121 CCTTACCTGCGCCAGGAG 180

DB 327 CCTTACCTGCGCCAGGAG 386  
QY 181 CCCCAGGCTCAGCAAG 240  
DB 387 CCCCAGGCTCAGCAAG 446  
QY 241 GGGAG 300  
DB 447 GGGAG 506  
QY 301 ACCCTGTGTCACCGGAG 351  
DB 507 ACCCTGTGTCACCGGAG 557

## RESULT 6

US-10-032-159A-17  
Sequence 17, Application US/10032159A  
Publication No. US20020164703A1  
GENERAL INFORMATION:  
APPLICANT: Pawlowski, Krzysztof  
APPLICANT: Reed, John C.  
TITLE OF INVENTION: CARD-DOMAIN CONTAINING POLYPEPTIDES,  
TITLE OF INVENTION: ENCODING NUCLEIC ACIDS, AND METHODS OF USE  
FILE REFERENCE: P-LJ 5100  
CURRENT APPLICATION NUMBER: US/10/032,159A  
CURRENT FILING DATE: 2001-12-19  
PRIOR APPLICATION NUMBER: US 60/257,457  
PRIOR FILING DATE: 2000-12-21  
NUMBER OF SEQ ID NOS: 37  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 17  
LENGTH: 276  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(276)  
US-10-032-159A-17

Query Match 66.2%; Score 276; DB 14; Length 276;  
Best Local Similarity 100.0%; Pred. No. 2.1e-68;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 GAG 105  
DB 1 GAG 60  
QY 106 CCCCAGGCTCAGCAAG 165  
DB 61 CCCCAGGCTCAGCAAG 120  
QY 166 GAGGTGTGCAAG 225  
DB 121 GAGGTGTGCAAG 180  
QY 226 TTGCTGAAG 285  
DB 181 TTGCTGAAG 240  
QY 286 AACCTGAGCTTACAG 321  
DB 241 AACCTGAGCTTACAG 276

## RESULT 7

US-09-798-412-9  
Sequence 9, Application US/09798412  
Publication No. US20030109428A1  
GENERAL INFORMATION:  
APPLICANT: Bertin, John  
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED



QY 18 CAGGACTCCGCACTGACGCGCACTGAGACGAGACACTGTGGAGATGAGAGCA 77  
 DB 82 CGGGGCGGGCTCGGGGCTTGAAGCGAGAGAGCGCGCTGTGGAGCAATGAGGGCT 141  
 QY 78 CGCCCAAGAGATGTAAGCTGACATGCTCCCAAGCGCGCTCAACCCCTTACCTGGCCAGGC 137  
 DB 142 CGCGGATCGGCTGGCTGGCGCCCTGMAACCGGCGCAAGCTCAAGCCCTATCTGGCCAGTG 201  
 QY 138 CAGGTGCTGTGCACTGAGACGAGAGAGAGTGTGCAAGCGCCGCGCTGCAACAG 197  
 DB 202 CGCGGTATGACGAGACGAGAGAGAGGTGTGAGACCTACCGCTTCCCGCGCG 261  
 QY 198 CGCGATCGGGGCGGGGCACTTGTGCTGTAAGACTCGAGGAGAAAGAGGGGCGCT 257  
 DB 262 CGTCAACCGGACGCGGGGCTGTGAGACATCTTGCGCGCTGCGCAAGAGGGGCTATGA 321  
 QY 258 CGGCTTCTGAGAGCGCTGAGTTCACCAACCTGACGCTGACCGCTGTCACGGGGCT 317  
 DB 322 GGGCTTCTGAGAGCGCTTGAAGTTCATCAACCGGACATTCACGCTGCTACCGGGCA 381  
 QY 318 GCAGCTGATGTTGACTTCAGTACTTAACTTAAAGCGGTGAG 354  
 DB 382 GGAACCGCGCCAGCGCTGTCTCCATGATCTTCATGAG 418

## RESULT 10

US-10-325-917-7  
 ; Sequence 7, Application US/10325917  
 ; Publication No. US20030113787A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bertin, John  
 ; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED  
 ; FILE OF INVENTION: PROTEIN FAMILY AND USES THEREOF  
 ; FILE REFERENCE: 07334-327001  
 ; CURRENT APPLICATION NUMBER: US/10/325,917  
 ; CURRENT FILING DATE: 2002-12-20  
 ; PRIOR APPLICATION NUMBER: US/09/798,412  
 ; PRIOR FILING DATE: 2001-03-02  
 ; PRIOR APPLICATION NUMBER: US 09/728,260  
 ; PRIOR FILING DATE: 2000-12-01  
 ; PRIOR APPLICATION NUMBER: US 09/685,791  
 ; PRIOR FILING DATE: 2000-10-10  
 ; PRIOR APPLICATION NUMBER: US 09/513,904  
 ; PRIOR FILING DATE: 2000-02-25  
 ; PRIOR APPLICATION NUMBER: US 60/168,780  
 ; PRIOR FILING DATE: 1999-12-03  
 ; NUMBER OF SEQ ID NOS: 19  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 7  
 ; LENGTH: 3949  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURES:  
 ; NAME/KEY: CDS  
 ; LOCATION: (41)...(3136)  
 ; US-10-325-917-7

Query Match 28.2%; Score 117.8; DB 15; Length 3949;

Best Local Similarity 59.3%; Pred. No. 1.6e-23;

Matches 200; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

QY 18 CAGGACTCCGCACTGACGCGCACTGAGACGAGACACTGTGGAGATGAGAGCA 77  
 DB 82 CGGGGCGGGCTCGGGGCTTGAAGCGAGAGAGCGCGCTGTGGAGCAATGAGGGCT 141  
 QY 78 CGCGCAAGAGATGTAAGCTGACATGCTCCCAAGCGCGCTCAACCCCTTACCTGGCCAGGC 137  
 DB 142 CGCGGATCGGCTGGCTGGCGCCCTGMAACCGGCGCAAGCTCAAGCCCTATCTGGCCAGTG 201  
 QY 138 CAGGTGCTGTGCACTGAGACGAGAGAGAGTGTGCAAGCGCCGCGCTGCAACAG 197

DB 202 CGGGTATTCAGACGAGACGAGAGAGAGTGTGTAGACCTTACCGCTTCCGCTGCG 261  
 QY 198 CGCCATCGGGGCGGGGCACTTGTGATTTGTGAAGCTCGAGGAGAAAGAGGGGCGAT 257  
 DB 262 CGTCAACCGGACCGGGGCGCTGTGAGACATCTTGCGCGCTGCGCAAGAGGGGCTATGA 321  
 QY 258 CGGCTTCTGAGAGCGCTGAGTTCACCAACCTGACGCTTACACCTGTGTACCGGGCT 317  
 DB 322 GGGCTTCTGAGAGCGCTTGAAGTTCATCAACCGGACATTCACGCTGCTACCGGGCA 381  
 QY 318 GCAGCTGATGTTGACTTCAGTACTTAACTTAAAGCGGTGAG 354  
 DB 382 GGAACCGCGCCAGCGCTGTCTCCATGATCTTCATGAG 418

## RESULT 11

US-10-032-159A-37  
 ; Sequence 37, Application US/10032159A  
 ; Publication No. US20020164703A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Pawlowski, Krzysztof  
 ; APPLICANT: Reed, John C.  
 ; APPLICANT: Godzik, Adam  
 ; TITLE OF INVENTION: CARD-DOMAIN CONTAINING POLYPEPTIDES,  
 ; FILE OF INVENTION: ENCODING NUCLEIC ACIDS, AND METHODS OF USE  
 ; FILE REFERENCE: P-LJ 5100  
 ; CURRENT APPLICATION NUMBER: US/10/032,159A  
 ; CURRENT FILING DATE: 2001-12-19  
 ; PRIOR APPLICATION NUMBER: US 60/257,457  
 ; PRIOR FILING DATE: 2000-12-21  
 ; NUMBER OF SEQ ID NOS: 37  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 37  
 ; LENGTH: 1141  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-10-032-159A-37

Query Match 26.8%; Score 111.8; DB 14; Length 1141;

Best Local Similarity 94.3%; Pred. No. 6.7e-22;

Matches 116; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATGGGGAACTGTGCGGAGGAGCTCCGCACTGACGCGCACTGAGACGAGACACTGTGG 60  
 DB 161 ATGGGGAACTGTGCGGAGGAGCTCCGCACTGACGCGCACTGAGACGAGACACTGTGG 220  
 QY 61 GAGATGATGAGAGCCACCGGACAGATCGTACGCTGATCTGCCCGGCGCTCAC 120  
 DB 221 GAGATGATGAGAGCCACCGGACAGATCGTACGCTGATCTGCCCGGACAGATTA 280  
 QY 121 CCC 123  
 DB 281 CCC 283

## RESULT 12

US-09-798-412-12  
 ; Sequence 12, Application US/09798412  
 ; Publication No. US20030109428A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bertin, John  
 ; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED  
 ; FILE OF INVENTION: PROTEIN FAMILY AND USES THEREOF  
 ; FILE REFERENCE: 07334-327001  
 ; CURRENT APPLICATION NUMBER: US/09/798,412  
 ; CURRENT FILING DATE: 2001-03-02  
 ; PRIOR APPLICATION NUMBER: US 09/728,260  
 ; PRIOR FILING DATE: 2000-12-01  
 ; PRIOR APPLICATION NUMBER: US 09/685,791  
 ; PRIOR FILING DATE: 2000-10-10  
 ; PRIOR APPLICATION NUMBER: US 09/513,904  
 ; PRIOR FILING DATE: 2000-02-25



PRIOR APPLICATION NUMBER: US 09/507,533  
 PRIOR FILING DATE: 2000-02-18  
 PRIOR APPLICATION NUMBER: US 60/168,780  
 PRIOR FILING DATE: 1999-12-03  
 NUMBER OF SEQ ID NOS: 19  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO: 12  
 LENGTH: 3441  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 US-09-798-412-12

Query Match 24.9%; Score 104; DB 11; Length 3441;  
 Best Local Similarity 58.3%; Pred. No. 1.2e-19;  
 Matches 182; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 43 GACGAGAGACACTGTGGAGATGATGAGAGCCACCGGCACAGATGTAAGCTGCATC 102  
 DB 31 GAAGAGAGACGCTTGTGGAGATGATGAGATGTAAGCTGCATC 90  
 QY 103 TCCCCAGCGCCTTACCCCTTACCTGCGCAGCCCAAGTGTGTGCCAGCTGAGACGAG 162  
 DB 91 AACCTGCCAAGCTACAGCCCTTACCTGCGCTCACTGATGATGATGAGAGATGAA 150  
 QY 163 GAGAGAGTCTGACAGCCCGGCTCACCAACAGCGCCAGTGGGCGGCACTTGCTG 222  
 DB 151 GATGAAGTGTATATGCTTATGCTGCTCATCAAGATCAACGAGCAGCGCGCTGTTG 210  
 QY 223 GATTGCTGAAGACTGAGAGGAAAGAGGGGCAATGCTTCTTGAGAGCTGAAGTTC 282  
 DB 211 GACATTTCACTACCAAGGGGCAAGGGGCTATATGTGTCTTCTTGAGAGCTTAATTT 270  
 QY 283 CACAACCTGACCTTACACCTGCTCACCGGCTCCAGCTGATTTAGTTAGTTAC 342  
 DB 271 TATTACCAAGACTGTACAACTGTGACTGGGAAAGAGCCCACTCGAGATTCTCCACC 330  
 QY 343 TTAGCGGTGAG 354  
 DB 331 ATTGTGTGTGAG 342

## RESULT 13

US-10-325-917-12  
 Sequence 12, Application US/10325917  
 Publication No. US20030113787A1

GENERAL INFORMATION:  
 APPLICANT: Bertlin, John  
 TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED  
 TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF  
 FILE REFERENCE: 07334-327001

CURRENT APPLICATION NUMBER: US/10/325,917  
 CURRENT FILING DATE: 2002-12-20

PRIOR APPLICATION NUMBER: US/09/798,412

PRIOR FILING DATE: 2001-03-02

PRIOR APPLICATION NUMBER: US 09/728,260

PRIOR FILING DATE: 2000-12-01

PRIOR APPLICATION NUMBER: US 09/685,791

PRIOR FILING DATE: 2000-10-10

PRIOR APPLICATION NUMBER: US 09/513,904

PRIOR FILING DATE: 2000-02-25

PRIOR APPLICATION NUMBER: US 09/507,533

PRIOR FILING DATE: 2000-02-18

PRIOR APPLICATION NUMBER: US 60/168,780

PRIOR FILING DATE: 1999-12-03

NUMBER OF SEQ ID NOS: 19

SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO: 12  
 LENGTH: 3441  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 US-10-325-917-12

Query Match 24.9%; Score 104; DB 15; Length 3441;

Best Local Similarity 58.3%; Pred. No. 1.2e-19;  
 Matches 182; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 43 GACGAGAGACACTGTGGAGATGATGAGAGCCACCGGCACAGATGTAAGCTGCATC 102  
 DB 31 GAAGAGAGACGCTTGTGGAGATGATGAGATGTAAGCTGCATC 90  
 QY 103 TCCCCAGCGCCTTACCCCTTACCTGCGCAGCCCAAGTGTGTGCCAGCTGAGACGAG 162  
 DB 91 AACCTGCCAAGCTACAGCCCTTACCTGCGCTCACTGATGATGATGAGAGATGAA 150  
 QY 163 GAGAGAGTCTGACAGCCCGGCTCACCAACAGCGCCAGTGGGCGGCACTTGCTG 222  
 DB 151 GATGAAGTGTATATGCTTATGCTGCTCATCAAGATCAACGAGCAGCGCGCTGTTG 210  
 QY 223 GATTGCTGAAGACTGAGAGGAAAGAGGGGCAATGCTTCTTGAGAGCTGAAGTTC 282  
 DB 211 GACATTTCACTACCAAGGGGCAAGGGGCTATATGTGTCTTCTTGAGAGCTTAATTT 270  
 QY 283 CACAACCTGACCTTACACCTGCTCACCGGCTCCAGCTGATTTAGTTAGTTAC 342  
 DB 271 TATTACCAAGACTGTACAACTGTGACTGGGAAAGAGCCCACTCGAGATTCTCCACC 330  
 QY 343 TTAGCGGTGAG 354  
 DB 331 ATTGTGTGTGAG 342

## RESULT 14

US-10-032-159A-7  
 Sequence 7, Application US/10032159A  
 Publication No. US20020164703A1

GENERAL INFORMATION:

APPLICANT: Pawlowski, Krzysztof

APPLICANT: Reed, John C.

TITLE OF INVENTION: CARD-DOMAIN CONTAINING POLYPEPTIDES,  
 TITLE OF INVENTION: ENCODING NUCLEIC ACIDS, AND METHODS OF USE

FILE REFERENCE: P-1J 5100

CURRENT APPLICATION NUMBER: US/10/032,159A

CURRENT FILING DATE: 2001-12-19

PRIOR APPLICATION NUMBER: US 60/257,457

PRIOR FILING DATE: 2000-12-21

NUMBER OF SEQ ID NOS: 37

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO: 7

LENGTH: 3744

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (1)...(3744)

US-10-032-159A-7

Query Match 24.9%; Score 104; DB 14; Length 3744;  
 Best Local Similarity 58.3%; Pred. No. 1.2e-19;  
 Matches 182; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 43 GACGAGAGACACTGTGGAGATGATGAGAGCCACCGGCACAGATGTAAGCTGCATC 102  
 DB 31 GAAGAGAGACGCTTGTGGAGATGATGAGATGTAAGCTGCATC 90  
 QY 103 TCCCCAGCGCCTTACCCCTTACCTGCGCAGCCCAAGTGTGTGCCAGCTGAGACGAG 162  
 DB 91 AACCTGCCAAGCTACAGCCCTTACCTGCGCTCACTGATGATGATGAGAGATGAA 150  
 QY 163 GAGAGAGTCTGACAGCCCGGCTCACCAACAGCGCCAGTGGGCGGCACTTGCTG 222  
 DB 151 GATGAAGTGTATATGCTTATGCTGCTCATCAAGATCAACGAGCAGCGCGCTGTTG 210  
 QY 223 GATTGCTGAAGACTGAGAGGAAAGAGGGGCAATGCTTCTTGAGAGCTGAAGTTC 282  
 DB 211 GACATTTCACTACCAAGGGGCAAGGGGCTATATGTGTCTTCTTGAGAGCTTAATTT 270



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 18, 2004, 03:27:47 ; Search time 74 Seconds  
(without alignments)  
298.149 Million cell updates/sec

Title: US-10-032-159a-16  
Perfect score: 720  
Sequence: 1 MSBLCRDBALTRALDBETLM.....SDPDGAGTSNRLVLPX 139

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues  
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	719	99.9	139	24	ABG76062 Human caspase recr
2	719	99.9	139	22	AAE07165 Human predicted ca
3	617.5	85.8	1004	22	AAE07164 Human caspase recr
4	270.5	37.6	1147	22	AAU01207 Human caspase recr
5	270.5	37.6	1147	23	AAU01207 Human CARD-11. Ho
6	270.5	37.6	1247	24	ABG76061 Human caspase recr
7	246	34.2	1032	22	AAU01206 Human caspase recr
8	246	34.2	1032	23	AAU01206 Human CARD-10. Ho
9	246	34.2	1032	23	AAU73247 Human plakoglobin

10	223.5	31.0	536	22	AAU01204 Rat caspase recruit
11	223.5	31.0	536	22	AAU01204 Rat CARD-9. Ratu
12	215.5	29.9	366	22	AAE07165 Human protein sequ
13	215.5	29.9	366	24	ABG76060 Caspase recruitmen
14	215.5	29.9	536	22	AAU01205 Human caspase recr
15	215.5	29.9	536	23	AAU01205 Human CARD-9. Hom
16	209	29.0	174	21	AAE07165 Human CARD-9. Hom
17	209	29.0	174	23	AAE07165 Human CARD-9. Hom
18	131	18.2	1139	22	AAU01207 Human ORF831
19	100	13.9	237	23	AAU01207 Human CARD-9. Hom
20	83	11.5	519	21	AAE07165 Human CARD-9. Hom
21	80	11.1	519	21	AAE07165 Human CARD-9. Hom
22	79	11.0	519	21	AAE07165 Human CARD-9. Hom
23	79	11.0	519	21	AAE07165 Human CARD-9. Hom
24	79	11.0	519	21	AAE07165 Human CARD-9. Hom
25	79	11.0	519	21	AAE07165 Human CARD-9. Hom
26	79	11.0	519	21	AAE07165 Human CARD-9. Hom
27	78	10.8	233	21	AAE07165 Human CARD-9. Hom
28	74.5	10.3	269	22	ABG02857 Novel human diapo
29	72.5	10.1	706	23	ABP40063 Staphylococcus epi
30	72.5	10.1	845	22	ABE62651 Drosophila melanog
31	72	10.0	299	23	AAO14117 Human SCZ gene (sc
32	71.5	9.9	260	22	ABE02008 Human transcrip
33	71.5	9.9	660	23	ABE02008 Human transcrip
34	71.5	9.9	707	22	ABE02008 Human transcrip
35	71.5	9.9	889	23	AAE02068 Human lung specifi
36	71.5	9.9	956	22	AAE02068 Human lung specifi
37	71.5	9.9	956	23	ABE02068 Human lung specifi
38	70.5	9.8	1266	23	AAE02068 Human lung specifi
39	69.5	9.7	352	21	AAE07165 Human CARD-9. Hom
40	69.5	9.7	497	23	ABE02068 Human lung specifi
41	69	9.6	203	20	AAE07165 Human CARD-9. Hom
42	69	9.6	361	21	AAE07165 Human CARD-9. Hom
43	69	9.6	361	21	AAE07165 Human CARD-9. Hom
44	69	9.6	371	21	AAE07165 Human CARD-9. Hom
45	69	9.6	371	21	AAE07165 Human CARD-9. Hom

## ALIGNMENTS

RESULT 1  
ABG76062  
ID ABG76062 standard; protein. 139 AA.  
AC ABG76062;  
DT 09-MAY-2003 (first entry)  
DE Human caspase recruitment domain containing protein, CARD-12X.

CARD; caspase recruitment domain; apoptosis; cell adhesion; inflammation;  
Cytokine receptor signaling; cancer; glioma; carcinoma; adenocarcinoma;  
CARD-containing polypeptide associated disorder; sarcoma; melanoma;  
hamaroma; leukemia; lymphoma; keratinocyte hyperplasia; neoplasia;  
keloid; benign prostatic hyperplasia; inflammatory hyperplasia; fibrosis;  
restenosis; allergy; arthritis; lupus; Sjogren's syndrome; sepsis; human;  
Crohn's disease; ulcerative colitis; graft versus host disease; stroke;  
abnormal cell death disease; myocardial infarction; heart failure;  
neurodegenerative disease; Parkinson's disease; Alzheimer's disease; HIV;  
CARD-12X; caspase activator; caspase inhibitor.

Key Location/Qualifiers  
Domain 16.107  
/label= CARD  
/note= "Caspase recruitment domain. Specifically  
claimed in claim 12"  
Misc-difference 139  
/label= Unknown  
/note= "Encoded by GNN"

XX	PN	US62002164703-AA.
XX	PD	07-NOV-2002.
XX	PF	19-DEC-2001; 2001US-0032159.
XX	PR	21-DEC-2000; 2000US-257457P.
XX	PA	(PAML/) PAWLOWSKI K.
XX	PA	(REED/) REED J.C.
XX	PA	(GODZ/) GODZIK A.
XX	PI	Pawlowski K., Reed JC., Godzik A;
DR	N-PSDB;	ABX11431.
XX	NP1;	2003-286137/28.
PT	New isolated CARD-containing nucleic acids,	useful for the diagnosis
PT	and treatment of disorders with aberrant expression or activity of the	CARD-containing polypeptide, such as cancer, stroke, arthritis, heart
PT	failure and AIDS -	
XX	Claim 11; Fig 3; 34pp;	English.
CC	The invention relates to an isolated nucleic acid molecule encoding a	caspase recruitment domain (CARD) containing polypeptide. CARD containing
CC	polypeptides are involved in apoptosis (as caspase activators and caspase	inhibitors), cell adhesion, inflammation and cytokine receptor
CC	signalling. The methods and compositions of the present invention are	useful for the diagnosis and treatment of disorders associated with the
CC	aberrant expression or activity of the CARD containing polypeptide such	as cancer, glioma, carcinoma, adenocarcinoma, sarcoma, melanoma, keloid,
CC	hamartoma, leukaemia, lymphoma, keratinocyte hyperplasia, neoplasia,	benign prostatic hypertrophy, inflammatory diseases such as arthritis, lupus,
CC	restenosis, allergies, inflammatory diseases such as Crohn's disease,	Sjogren's syndrome, ulcerative colitis, graft versus
CC	host disease, sepsis, abnormal cell death diseases such as stroke,	myocardial infarction, heart failure, neurodegenerative diseases like
CC	Parkinson's disease and Alzheimer's disease, and HIV infection. The	present sequence represents the amino acid sequence of the human
CC	caspase recruitment domain containing protein, CARD-12X.	
XX	Sequence	139 AA;
SQ	Query Match	99.9%; Score 719; DB 24; Length 139;
	Best Local Similarity	100.0%; Pred. No. 4,18-81;
	Matches 138; Conservative	0; Mismatches 0; Indels 0; Gaps 0
OY	1 MGELCRDSDALTLMDDEITLMMEMSHRRRIIVACIPSPILTYLRQAAYLCQLDEEYVLHS	60
Db	1 MGELCRDSDALTLMDDEITLMMEMSHRRRIIVACIPSPILTYLRQAAYLCQLDEEYVLHS	60
OY	61 PRLTNSAMRGHLLDLKTRGNKAIAFESLKFNPPVTLYLVGLDPDVDFSNFSSESS	120
Db	61 PRLTNSAMRGHLLDLKTRGNKAIAFESLKFNPPVTLYLVGLDPDVDFSNFSSESS	120
OY	121 DFDGLAGTSNRLRLVLP	138
Db	121 DFDGLAGTSNRLRLVLP	138
RESULT 2		
ID	AAE07165 standard; Protein; 1139 AA.	
AC	AAE07165;	
DT	06-NOV-2001 (first entry)	
XX	Human predicted caspase recruitment domain-14 (CARD-14).	
XX	Human, caspase recruitment domain-14; CARD-14; chromosome 17;	
KM	nuclear factor-kappa B; NF-kB; cell growth; cell death; cancer; therapy;	

OS	Homo sapiens.
XX	
FH	Key
FT	Misc-difference
FT	700
XX	/note= "Encoded by TGG"
PN	W0200159065-A2.
XX	
PD	16-AUG-2001.
XX	
PF	22-JAN-2001; 2001WO-US02087.
XX	
PR	09-FEB-2000; 2000US-0181159.
XX	
PA	(MILL-) MILLENNITUM PHARM INC.
PI	Bertin J;
XX	
DR	WPJ: 2001-497073/54.
XX	
DR	N-PSDB; AADJ3448.
PT	An isolated caspase recruitment domain polypeptide useful for
PT	regulating growth and cell death and useful for the treatment of cancer
XX	-
PS	Disclosure; Fig 2a-2c; 109pp; English.
XX	
CC	The present sequence is predicted human caspase recruitment domain-14
CC	(CARD-14). The CARD-14 gene is located on chromosome 17. The CARD-14 is
CC	used for the detection of modulators that modulates the ability of
CC	CARD-14 to bind to Bcl-10 and stimulates phosphorylation of Bcl-10 or
CC	activation of nuclear factor-kappa B (NF-kB). The CARD-14 is useful for
CC	regulating growth and cell death and useful for the treatment of cancer.
CC	It is also useful for the treatment of autoimmune disorders (e.g.,
CC	systemic lupus erythematosus), neurological disorders e.g., Alzheimer's
CC	disease, Parkinson's disease, inflammatory disorders, haematological disorders
CC	(e.g., anaemia, myelodysplastic syndromes), myocardial infarctions,
CC	strokes, immune disorders (e.g., Crohn's disease, allergic rhinitis),
CC	cell signalling disorders and certain viral and bacterial infections.
XX	
SQ	Sequence 1139 AA;
	Query Match 99.9%; Score 719; DB 22; Length 1139;
	Best Local Similarity 100.0%; Pred. No. 9.5e-81;
	Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0
OY	1 MGELCRDSDALNALDEETLWMESHRRIRYCICPSRLTPPYRQAVALCOLDEEFTLS 60
DB	1 MGELCRDSDALNALDEETLWMESHRRIRYCICPSRLTPPYRQAVALCOLDEEFTLS 60
OY	61 PRLTNSAMAGHLIDLKTRGNCAIAFLSILKFHPDPVTLVTGLQPDVPFNFSGESS 120
DB	61 PRLTNSAMAGHLIDLKTRGNCAIAFLSILKFHPDPVTLVTGLQPDVPFNFSGESS 120
OY	121 DFDGLAGTSRNRLIIVTP 138
DB	121 DFDGLAGTSRNRLIIVTP 138
RESULT_3	
ID	AAB07164 standard; Protein; 1004 AA.
NC	AAB07164;
XX	
DT	06-NOV-2001 (first entry)

XX Human caspase recruitment domain-14 (CARD-14).  
XX Human, caspase recruitment domain-14; CARD-14; chromosome 17;  
XX nuclear factor-kappa B; NF-kB; cell growth; cell death; cancer; therapy;  
KW autoimmune disorder; systemic lupus erythematosus; neurological disorder;  
KW Alzheimer's disease; Parkinson's disease; inflammatory disorder; anaemia;  
KW haematological disorder; myelodysplastic syndrome; myocardial infarction;  
KW stroke; immune disorder; Crohn's disease; allergic rhinitis; infection;  
KW cell signalling disorder; cytostatic; immunosuppressive; noctropic;  
KW neuroprotective; antiviral; antibacterial.  
XX Homo sapiens.  
XX  
XX Key  
FT Modified-site  
FT Location/Qualifiers  
FT /note= "cAMP- and cGMP-dependent protein kinase  
FT phosphorylation site"  
FT 10..116  
FT Domain  
FT /label= CARD\_domain  
FT 12..15  
FT /note= "Casein kinase II phosphorylation site"  
FT 18..21  
FT /note= "Casein kinase II phosphorylation site"  
FT 25..27  
FT /note= "Protein kinase C phosphorylation site"  
FT 60..62  
FT /note= "Protein kinase C phosphorylation site"  
FT 91..93  
FT /note= "Protein kinase C phosphorylation site"  
FT 114..117  
FT /note= "Protein kinase C phosphorylation site"  
FT 117..122  
FT /note= "N-myristoylation site"  
FT 121..123  
FT /note= "Protein kinase C phosphorylation site"  
FT 126..420  
FT Domain  
FT /label= Coiled\_Coil\_domain  
FT 130..135  
FT /note= "N-myristoylation site"  
FT 134..137  
FT /note= "Casein kinase II phosphorylation site"  
FT 161..166  
FT /note= "N-myristoylation site"  
FT 165..168  
FT /note= "Casein kinase II phosphorylation site"  
FT 220..227  
FT /note= "Tyrosine kinase phosphorylation site"  
FT 221..224  
FT /note= "Casein kinase II phosphorylation site"  
FT 239..325  
FT /label= K-Box\_domain  
FT 240..243  
FT /note= "Casein kinase II phosphorylation site"  
FT 250..252  
FT /note= "Protein kinase C phosphorylation site"  
FT 253..256  
FT /note= "Casein kinase II phosphorylation site"  
FT 259..262  
FT /note= "Casein kinase II phosphorylation site"  
FT 280..283  
FT /note= "Casein kinase II phosphorylation site"  
FT 290..293  
FT /note= "Casein kinase II phosphorylation site"  
FT 297..300  
FT /note= "Casein kinase II phosphorylation site"  
FT 307..309  
FT /note= "Protein kinase C phosphorylation site"  
FT 307..310  
FT /note= "Casein kinase II phosphorylation site"  
FT 359..365  
FT /note= "Tyrosine kinase phosphorylation site"  
FT 366..368  
FT Modified-site  
FT /note= "Protein kinase C phosphorylation site"  
FT 366..369  
FT /note= "Casein kinase II phosphorylation site"  
FT 378..381  
FT /note= "Casein kinase II phosphorylation site"  
FT 384..386  
FT /note= "Protein kinase C phosphorylation site"  
FT 385..406  
FT /note= "Leucine zipper pattern"  
FT 449..452  
FT /note= "Casein kinase II phosphorylation site"  
FT 463..466  
FT /note= "Casein kinase II phosphorylation site"  
FT 463..465  
FT /note= "Protein kinase C phosphorylation site"  
FT 470..472  
FT /note= "Protein kinase C phosphorylation site"  
FT 501..504  
FT /note= "Casein kinase II phosphorylation site"  
FT 511..516  
FT /note= "N-myristoylation site"  
FT 568..660  
FT Domain  
FT /label= PDZ\_domain  
FT 587..592  
FT /note= "N-myristoylation site"  
FT 589..592  
FT /note= "N-glycosylation site"  
FT 602..605  
FT /note= "Casein kinase II phosphorylation site"  
FT 634..637  
FT /note= "Casein kinase II phosphorylation site"  
FT 653..655  
FT /note= "Protein kinase C phosphorylation site"  
FT 674..677  
FT /note= "Casein kinase II phosphorylation site"  
FT 676..745  
FT Domain  
FT /label= SH3\_domain  
FT 714..719  
FT /note= "N-myristoylation site"  
FT 725..727  
FT /note= "Protein kinase C phosphorylation site"  
FT 725..728  
FT /note= "Casein kinase II phosphorylation site"  
FT 733..738  
FT /note= "N-myristoylation site"  
FT 737..740  
FT /note= "N-glycosylation site"  
FT 759..761  
FT /note= "Protein kinase C phosphorylation site"  
FT 760..763  
FT /note= "cAMP- and cGMP-dependent protein kinase  
FT phosphorylation site"  
FT 785..793  
FT /note= "Peroxisomal targeting signal"  
FT 796..799  
FT /note= "Casein kinase II phosphorylation site"  
FT 800..805  
FT /note= "N-myristoylation site"  
FT 826..1004  
FT Domain  
FT /label= Guanylate\_kinase\_domain  
FT 842..844  
FT /note= "Protein kinase C phosphorylation site"  
FT 860..863  
FT /note= "Casein kinase II phosphorylation site"  
FT 868..870  
FT /note= "Protein kinase C phosphorylation site"  
FT 870..872  
FT /note= "RGD cell attachment sequence"  
FT 893..896  
FT /note= "Casein kinase II phosphorylation site"  
FT 926..929  
FT /note= "Casein kinase II phosphorylation site"  
FT 941..949  
FT Peptide



```

FT Domain 635..748 /note="PDZ domain"
FT Domain 635..1147 /note="MAGUK domain"
FT Modified-site 638..641 /note="MAGUK domain"
FT Modified-site 678..683 /note="Glycosaminoglycan attachment site"
FT Modified-site 687..689 /note="N-myristoylation site"
FT Modified-site 692..695 /note="Protein kinase C phosphorylation site"
FT Modified-site 698..703 /note="Casein kinase II phosphorylation site"
FT Modified-site 710..715 /note="N-myristoylation site"
FT Modified-site 725..728 /note="N-myristoylation site"
FT Modified-site 761..766 /note="Casein kinase II phosphorylation site"
FT Modified-site 764..767 /note="N-myristoylation site"
FT Modified-site 766..834 /note="Casein kinase II phosphorylation site"
FT Domain 776..779 /note="SH3 domain"
FT Modified-site 779..782 /note="N-glycosylation site"
FT Modified-site 787..789 /note="Casein kinase II phosphorylation site"
FT Modified-site 816..819 /note="Protein kinase C phosphorylation site"
FT Modified-site 823..828 /note="Casein kinase II phosphorylation site"
FT Modified-site 847..850 /note="N-myristoylation site"
FT Modified-site 853..858 /note="Casein kinase II phosphorylation site"
FT Modified-site 857..859 /note="N-myristoylation site"
FT Modified-site 872..875 /note="Protein kinase C phosphorylation site"
FT Modified-site 882..1147 /note="Casein kinase II phosphorylation site"
FT Domain 897..900 /note="Guanylate kinase (GUK) domain"
FT Modified-site 917..922 /note="Casein kinase II phosphorylation site"
FT Modified-site 926..929 /note="N-myristoylation site"
FT Modified-site 935..937 /note="Casein kinase II phosphorylation site"
FT Modified-site 1003..1006 /note="Protein kinase C phosphorylation site"
FT Modified-site 1010..1018 /note="Casein kinase II phosphorylation site"
FT Modified-site 1050..1055 /note="Tyrosine kinase phosphorylation site"
FT Modified-site 1088..1091 /note="N-myristoylation site"
FT Modified-site 1120..1123 /note="Casein kinase II phosphorylation site"
FT Modified-site 1120..1123 /note="Casein kinase II phosphorylation site"

```

```

XX (MILL-) MILLENNIUM PHARM INC.
PA Bertin J;
XX WPI: 2001-367809/38.
XX DR N-PEDB; AAS05389.
XX
PT Novel caspase recruitment domain (CARD) proteins, CARD-9, CARD-10,
PT CARD-11, useful as targets for therapy, as immunogens, and in screening
PT and detection assays -
XX
PS Claim 9; Fig 14A-14C; 145pp; English.
XX
XX The present sequence represents novel human caspase recruitment
XX domain, CARD-11. The polynucleotide encoding this sequence was
XX isolated from a human T-cell cDNA library. Also described are
XX novel human sequences for CARD-9 and CARD-10 (AAU01205, AAU01206)
XX and rat CARD-9 (AAU01204). CARD-9, CARD-10 and CARD-11 interact
XX with Bcl-10 which is thought to activate nuclear factor (NF)-kappaB
XX and apoptosis. The sequences of the invention can be used for
XX treating a disorder associated with abnormal levels of apoptosis by
XX modulating the expression or activity of CARD-9, CARD-10, or CARD-11.
XX They can be used for the treatment of hyperproliferative disorders
XX (e.g. cancer), autoimmune disorders (e.g. systemic lupus erythematosus),
XX neurological disorders (e.g. Alzheimer's disease), inflammatory disorders
XX (e.g. Crohn's disease), and viral infection (e.g. HIV). The CARD
XX polypeptide, polynucleotide and an antibody which selectively binds to
XX CARD can be used in screening and detection assays (e.g. chromosomal
XX mapping, tissue typing), predictive medicine (prognostic assays),
XX monitoring clinical trials, and therapy (treatment and prophylaxis). The
XX CARD polypeptide may be used to screen for drugs that bind to and/or
XX modulate it. CARD sequences are potential targets for regulating
XX inflammation, cancer, NF-kappaB signaling, stress-related response and
XX apoptosis in human disease. A host cell containing a polynucleotide
XX encoding CARD can be used to create transgenic animals.
XX
SQ Sequence 1147 AA;
Query Match 37.6%; Score 270.5; DB 22; Length 1147;
Best Local Similarity 48.6%; Pred. No. 2.9e-24;
Matches 54; Conservative 23; Mismatches 33; Indels 1; Gaps 1;
QY 15 DEETMEMESHRRIRVRCIPSLTPYLRQAKLVCQDEBEVHSPRLTMSARAGTL 74
DB 11 EEDALMEVENCRRMLSRYPINPAKLTPLYRQCKVIDQDEBEVHAPVLPKINRAGLT 70
QY 75 DLKTRGNKAIAFLSLKFNPDVYTLVTGLQPDVPSNFGSSPDGL 125
DB 71 DILHTKGQRYVFLSLFLFYPELYKLVTKETRRPSTIVZEG-HEGL 120
RESULT 5
AAG79555 standard; Protein, 1147 AA.
AC AAG79555;
XX
XX 09-DEC-2002 (first entry)
XX
XX Human CARD-11.
XX
XX Rac; human; caspase recruitment domain; CARD-9; CARD-10;
XX CARD-11; apoptosis; inflammation; cell growth; cell death;
XX lymphocyte activation; cancer; melanoma; autoimmune disease;
XX arthritis; neurological disorder; Alzheimer's disease.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Domain 6..112 /label= CARD_domain
XX
XX Domain 130..431

```



```

FT      Domain /label= Coiled-coil_domain
FT      635..748
FT      /label= PDZ_domain
FT      766..834
FT      Domain /label= SH3_domain
FT      882..1147
FT      Domain /label= Guanylate_kinase_domain
FT      635..1147
FT      /label= MAGUK_domain
FT      635..1147
PN      WO200270652-A2.
PD      12-SEP-2002.
PF      28-FEB-2002; 2002WO-US06147.
PR      02-MAR-2001; 2001US-0798412.
PA      (MILL-) MILLENNIUM PHARM INC.
XX      Bertin J;
XX      WPI: 2002-698749/75.
XX      N-PSDB; ABA00334.
XX      CARD-9, CARD-10 or CARD-11 polypeptides and polynucleotides, useful for
XX      treating disorders associated with inappropriate apoptosis or
XX      lymphocyte activation, e.g. cancer
XX      Disclosure; Fig 14; 151pp; English.
XX      This sequence represents human caspase recruitment domain (CARD)-11.
XX      CARD proteins play roles in apoptotic and inflammatory signalling
XX      pathways. CARD-9, -10 and -11 participate in the network of
XX      interactions that modulate caspase activity. They are thought to be
XX      useful as modulating agents for regulating a variety of cellular
XX      processes including cell growth and cell death. CARD proteins and
XX      nucleic acids are useful for treating a disorder associated with
XX      inappropriate apoptosis or lymphocyte activation or for diagnosing
XX      subjects having or that are at risk of developing a disorder associated
XX      with aberrant CARD-9, CARD-10 or CARD-11 expression or activity, such
XX      as cancer e.g. melanoma, autoimmune disorders e.g. arthritis, or
XX      neurological disorders e.g. Alzheimer's disease.
XX      Sequence 1147 AA;
SQ
Query Match 37.6%; Score 270.5; DB 23; Length 1147;
Best Local Similarity 48.6%; Pred. No. 2.9e-24;
Matches 54; Conservative 23; Mismatches 33; Indels 1; Gaps 1;
QY      15 DEETIMWESHRRIRVRCICPSRLTPYLQAKVLCQDDEEVLHSPRLTNSAMRAGHL 74
DB      11 EEDALWENVECRHMLSRYNPAKLTPYLQCKVIDEDDEVANAPLPSKINRAGSL 70
QY      75 DLTKRGKNGAIAFBSLKFANPDVYTLVTCQDPVDSNBSGSSDDGL 125
DB      71 DLTHTKGGRGVVFLSEFYPYBELYKVTGKEPFRSTIWEHG-HEGL 120
RESULT 6
ID      ABG76061 standard; protein; 1247 AA.
AC      ABG76061;
XX
DT      09-MAY-2003 (first entry)
XX
DE      Human caspase recruitment domain containing protein, CARD-11X.
XX
KW      CARD; caspase recruitment domain; apoptosis; cell adhesion; inflammation;
KW      cytokine receptor signalling; cancer; glioma; carcinoma; adenocarcinoma;
KW      CARD-containing polypeptide associated disorder; sarcoma; melanoma;
KW      hamartoma; leukaemia; lymphoma; keratinocyte hyperplasia; neoplasia;

```

```

KW      keloid; benign prostatic hypertrophy; inflammatory hyperplasia; fibrosis;
KW      restenosis; allergy; arthritis; lupus; Sjogren's syndrome; sepsis; human;
KW      Crohn's disease; ulcerative colitis; graft versus host disease; stroke;
KW      abnormal cell death disease; myocardial infarction; heart failure;
KW      neurodegenerative disease; Parkinson's disease; Alzheimer's disease; HIV;
KW      CARD-11X; caspase activator; caspase inhibitor.
XX      Homo sapiens.
XX
FH      Key Location/Qualifiers
FT      Domain 12..103
FT      /label= CARD
FT      /note= "Caspase recruitment domain. Specifically
FT      claimed in claim 12"
FT      Domain 368..1344
FT      /label= ERM domain
FT      /note= "Erin/radixin/moesin domain. Specifically
FT      claimed in claim 12"
FT      Domain 2175..2514
FT      /label= PDZ domain
FT      /note= "Post synaptic density disc-large zo-1 domain.
FT      Specifically claimed in claim 12"
XX      US2002164703-A1.
XX      07-NOV-2002.
XX      19-DEC-2001; 2001US-0032159.
XX      21-DEC-2000; 2000US-257457P.
XX      (PAWL/) PAWLOWSKI K.
XX      (REED/) REED J C.
XX      (GODZ/) GODZIK A.
XX      Pawlowski K, Reed JC, Godzik A;
XX      WPI: 2003-288137/28.
XX      N-PSDB; ABX11430.
XX      New isolated CARD-containing nucleic acid, useful for the diagnosis
XX      and treatment of disorders with aberrant expression or activity of the
XX      CARD-containing polypeptide, such as cancer, stroke, arthritis, heart
XX      failure and AIDS
XX      Claim 11; Fig 2; 34pp; English.
XX
CC      The invention relates to an isolated nucleic acid molecule encoding a
CC      caspase recruitment domain (CARD) containing polypeptide. CARD containing
CC      polypeptides are involved in apoptosis (as caspase activators and caspase
CC      inhibitors), cell adhesion, inflammation and cytokine receptor
CC      signalling. The methods and compositions of the present invention are
CC      useful for the diagnosis and treatment of disorders associated with the
CC      aberrant expression or activity of the CARD containing polypeptide such
CC      as cancer, glioma, carcinoma, adenocarcinoma, sarcoma, melanoma, keloid,
CC      hamartoma, leukaemia, lymphoma, keratinocyte hyperplasia, neoplasia,
CC      benign prostatic hypertrophy, inflammatory hyperplasia, fibrosis,
CC      restenosis, allergies, inflammatory diseases such as arthritis, lupus,
CC      Sjogren's syndrome, Crohn's disease, ulcerative colitis, graft versus
CC      host disease, sepsis, abnormal cell death diseases such as stroke,
CC      myocardial infarction, heart failure, neurodegenerative diseases like
CC      Parkinson's disease and Alzheimer's disease, and HIV infection. The
CC      present sequence represents the amino acid sequence of the caspase
CC      recruitment domain containing protein, CARD-11X.
XX
SQ      Sequence 1247 AA;
Query Match 37.6%; Score 270.5; DB 24; Length 1247;
Best Local Similarity 48.6%; Pred. No. 3.3e-24;
Matches 54; Conservative 23; Mismatches 33; Indels 1; Gaps 1;
QY      15 DEETIMWESHRRIRVRCICPSRLTPYLQAKVLCQDDEEVLHSPRLTNSAMRAGHL 74

```

Db 11 EEDALWENECNRHMLSRYPINPAKLTPYLRCKQVIDEODEEVLNAPMLPSKINAGRLL 70  
QY 75 DLKTRGKGAIAFLESLKFNPDVYTLVGTGQPDVDFNFGESDPRGL 125  
Db 71 DILHTKGQGVVFLFSLFEPYLPYLYKLVTKGEPTRFSTIVEG-HGGL 120

RESULT 7  
AAU01206  
ID AAU01206 standard; Protein; 1032 AA.  
XX AAU01206;  
XX  
XX 12-SEP-2001 (first entry)  
DE Human caspase recruitment domain, CARD-10 polypeptide.  
XX  
XX Human; caspase recruitment domain; CARD-10; Bcl-10; NF-kappaB;  
KM apoptosis; hyperproliferative disorder; autoimmune; neurological;  
KM inflammatory disorder; viral infection; stress-related response.  
XX  
XX Homo sapiens.  
OS

Key Location/Qualifiers  
FH Modified-site 15..20 /note= "N-myristoylation site"  
FT Modified-site 18..21 /note= "Casein kinase II phosphorylation site"  
FT Domain 23..123 /note= "CARD domain"  
FT Modified-site 68..70 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 76..79 /note= "N-glycosylation site"  
FT Modified-site 78..80 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 88..91 /note= "Amidation site"  
FT Modified-site 112..115 /note= "Casein kinase II phosphorylation site"  
FT Modified-site 113..118 /note= "N-myristoylation site"  
FT Domain 147..457 /note= "Coiled coil domain"  
FT Modified-site 201..207 /note= "Tyrosine kinase phosphorylation site"  
FT Region 230..251 /note= "Leucine zipper homology region"  
FT Modified-site 242..245 /note= "Casein kinase II phosphorylation site"  
FT Modified-site 293..295 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 293..296 /note= "Casein kinase II phosphorylation site"  
FT Modified-site 309..314 /note= "N-myristoylation site"  
FT Modified-site 313..315 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 331..334 /note= "Casein kinase II phosphorylation site"  
FT Domain 366..398 /note= "Tropomyosin domain"  
FT Modified-site 412..415 /note= "Casein kinase II phosphorylation site"  
FT Region 426..447 /note= "Leucine zipper homology region"  
FT Modified-site 438..441 /note= "Casein kinase II phosphorylation site"  
FT Domain 457..1032 /note= "N-myristoylation site"  
FT Modified-site 472..475 /note= "N-glycosylation site"  
FT Modified-site 478..481

FT Modified-site /note= "Casein kinase II phosphorylation site"  
FT /note= "N-myristoylation site"  
FT Modified-site 510..513 /note= "Casein kinase II phosphorylation site"  
FT Modified-site 512..514 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 549..552 /note= "Casein kinase II phosphorylation site"  
FT Modified-site 558..560 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 565..570 /note= "N-myristoylation site"  
FT Modified-site 570..573 /note= "Casein kinase II phosphorylation site"  
FT Modified-site 595..598 /note= "N-glycosylation site"  
FT Modified-site 603..605 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 638..641 /note= "Glycosaminoglycan attachment site"  
FT Modified-site 642..644 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 656..661 /note= "N-myristoylation site"  
FT Modified-site 681..684 /note= "Casein kinase II phosphorylation site"  
FT Modified-site 690..693 /note= "Casein kinase II phosphorylation site"  
FT Domain 704..772 /note= "SH3 domain"  
FT Modified-site 712..715 /note= "N-glycosylation site"  
FT Modified-site 714..717 /note= "Casein kinase II phosphorylation site"  
FT Modified-site 733..739 /note= "Tyrosine kinase phosphorylation site"  
FT Modified-site 748..751 /note= "Casein kinase II phosphorylation site"  
FT Modified-site 754..756 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 754..757 /note= "Casein kinase II phosphorylation site"  
FT Modified-site 761..766 /note= "N-myristoylation site"  
FT Modified-site 782..784 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 809..814 /note= "N-myristoylation site"  
FT Domain 830..1032 /note= "Guanylate kinase (GUK) domain"  
FT Modified-site 830..832 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 868..870 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 869..872 /note= "Casein kinase II phosphorylation site"  
FT Modified-site 882..885 /note= "Casein kinase II phosphorylation site"  
FT Modified-site 893..898 /note= "N-myristoylation site"  
FT Modified-site 915..918 /note= "Amidation site"  
FT Modified-site 947..949 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 981..986 /note= "N-myristoylation site"  
FT Modified-site 1021..1026 /note= "N-glycosylation site"  
FT Modified-site 1022..1024 /note= "Protein kinase C phosphorylation site"  
FT Modified-site 1028..1031 /note= "Casein kinase II phosphorylation site"

FN	XX	W0200140668-A2.
PD	XX	07-JUN-2001.
PF	XX	01-DEC-2000; 2000MO-US32716.
PR	XX	03-DEC-1999; 99US-0168780.
PR	XX	18-FEB-2000; 2000US-0507533.
PR	XX	25-FEB-2000; 2000US-0513904.
PR	XX	10-OCT-2000; 2000US-0685791.
PA	XX	(MILL-) MILLENNIUM PHARM INC.
PT	XX	Bertin J;
DR	XX	WPI; 2001-367809/38.
XX	XX	N-PSDB; AAS05388.
PT	XX	Novel caspase recruitment domain (CARD) proteins, CARD-9, CARD-10,
PT	XX	CARD-11, useful as targets for therapy, as immunogens, and in screening
PT	XX	and detection assays -
PS	XX	Claim 9; Fig 10A-10C; 145pp; English.
CC	XX	The present sequence represents novel human caspase recruitment
CC	XX	domain, CARD-10. The polynucleotide encoding this sequence was isolated
CC	XX	from a human skin cDNA library. Also described are novel human sequences
CC	XX	for CARD-9 and CARD-11 (AAU01205, AAU01207) and rat CARD-9 (AAU01204).
CC	XX	CARD-9, CARD-10 and CARD-11 interact with Bcl-10 which is thought to
CC	XX	activate nuclear factor (NF)-kappaB and apoptosis. The sequences of the
CC	XX	invention can be used for treating a disorder associated with abnormal
CC	XX	levels of apoptosis by modulating the expression or activity of CARD-9,
CC	XX	CARD-10, or CARD-11. They can be used for the treatment of
CC	XX	hyperproliferative disorders (e.g. cancer), autoimmune disorders (e.g.
CC	XX	systemic lupus erythematosus), neurological disorders (e.g.
CC	XX	Alzheimer's disease), inflammatory disorders (e.g. Crohn's disease),
CC	XX	and viral infection (e.g. HIV). The CARD polypeptide, polynucleotide
CC	XX	and an antibody which selectively binds to CARD can be used in screening
CC	XX	and detection assays (e.g. chromosomal mapping, tissue typing),
CC	XX	predictive medicine (prognostic assays, monitoring clinical trials, and
CC	XX	therapy (treatment and prophylaxis). The CARD polypeptide may be used
CC	XX	to screen for drugs that bind to and/or modulate it. CARD sequences are
CC	XX	potential targets for regulating inflammation, cancer, NF-kappaB
CC	XX	signaling, stress-related response and apoptosis in human disease. A
CC	XX	host cell containing a polynucleotide encoding CARD can be used to
CC	XX	create transgenic animals.
SQ	XX	Sequence 1032 AA;
Query Match	34.2%; Score 246; DB 22; Length 1032;	
Best Local Similarity	48.9%; Pred. 3e-21;	
Matches	46; Conservative 21; Mismatches 27; Indels 0; Gaps	
QY	15 DEETLWEMESHRRIRVRCISRLTPYLRQAKVLCQLDDEEVLHSPRLTNSAMRAGHLL 74	
DB	23 EEDLWLRIGEVHRLRALNPAKLTPYLRQCRVIDEQQEVEEVLSTYRPPCRVNRTRGLM 82	
QY	75 DLKTRGKNGAIAFLBSLKFNPDVTVTVGLQAP 108	
DB	83 DILCRGKRGVEAFLEALERYYPHFHTLLTGQEP 116	
RESULT 8		
AAAG79554	standard; Protein; 1032 AA.	
AAAG79554;		
09-DEC-2002	(first entry)	
Human CARD-10.		

KW Ref.: human; caspase recruitment domain; CARD-9; CARD-10;  
KW CARD-11; apoptosis; inflammation; cell growth; cell death;  
KW lymphocyte activation; cancer; melanoma; autoimmune disease,  
arthritis; neurological disorder; Alzheimer's disease.  
XX Homo sapiens.

XN Key Location/Qualifiers  
FT Domain 23..123 /label= CARD\_domain  
FT FT 147..457 /label= coll-ed-coil\_domain  
FT Domain 704..772 /label= SH3\_domain  
FT FT /label= SH3\_domain  
FT Domain 830..1032 /label= Guanylate\_kinase\_domain  
FT FT 366..398 /label= Tropomyosin\_domain  
FT Domain /label= MAGUK\_domain

FN WO200270652-A2.  
PD 12-SEP-2002.  
PF 28-FEB-2002; 2002WO-USO6147.  
PR 02-MAR-2001; 2001US-0798412.  
PA (MILL-) MILLENNIUM PHARM INC.  
PI Bertin J;

DH MPI; 2002-698749/75.  
DR N-PSTB; ABA00333.

PT CARD-9, CARD-10 or CARD-11 polypeptides and polymucleotides, useful for treating disorders associated with inappropriate apoptosis or lymphocyte activation, e.g. cancer -  
PS Claim 14; Fig 10; 15pp; English.

CC This sequence represents human caspase recruitment domain (CARD)-10. CARD proteins play roles in apoptotic and inflammatory signalling pathways. CARD-9, -10 and -11 participate in the network of interactions that modulate caspase activity. They are thought to be useful as modulating agents for regulating a variety of cellular processes including cell growth and cell death. CARD proteins and nucleic acids are useful for treating a disorder associated with inappropriate apoptosis or lymphocyte activation or for diagnosing subjects having or that are at risk of developing a disorder associated with aberrant CARD-9, CARD-10 or CARD-11 expression or activity, such as cancer e.g. melanoma, autoimmune disorders e.g. arthritis, or neurological disorders e.g. Alzheimer's disease.

SQ Sequence 1032 AA:

Query Match 34.2%; Score 246; DB 23; Length 1032;  
Best Local Similarity 48.9%; Pred. No. 3e-21;  
Matches 46; Conservative 21; Mismatches 27; Indels 0; Gaps 0;

OY 15 DEETLMESESHRRIVRCICPSRLTPPYLAQAATLCOLDEEVLHSPLTNSAMRAGHL 74  
Db 23 EEDTLMERIRIGVNRRLRALNPAKLTFFYLACQGVIBDQEDEVLYSTRPFCVRVRTSLM 82

OY 75 DLKTRGNGALAFLESFKFANFDVTITLVGLDP 108  
Db 83 DILRGCKRGVEAFLDALBEFYEPHFLLTGOSP 116





XX 26-JUN-2001 (first entry)  
 XX Human protein sequence SEQ ID NO:18328.  
 DE Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
 XX Homo sapiens.  
 OS Homo sapiens.  
 XX EPI074617-A2.  
 XX 07-FEB-2001.  
 PD 28-JUL-2000; 2000EP-0116126.  
 PF 29-JUL-1999; 99JP-0248036.  
 XX 27-AUG-1999; 99JP-0300253.  
 PR 11-JAN-2000; 2000JP-0118776.  
 PR 02-MAY-2000; 2000JP-0183767.  
 PR 09-JUN-2000; 2000JP-0241899.  
 XX (HELI-) HELIX RES INST.  
 PA Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J,  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T,  
 XX WPI; 2001-318749/34.  
 DR Primer sets for synthesizing polynucleotides, particularly the 5602  
 XX full-length CDNA defined in the specification, and for the detection  
 PT full-length CDNA defined in the specification, and for the detection  
 PT full-length CDNA -  
 PS Claim 8; SEQ ID 18328; 2537bp + CD ROM; English.  
 XX The present invention describes primer sets for synthesising 5602  
 CC full-length CDNA defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesising polynucleotides,  
 CC particularly full-length CDNA. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length CDNA. The primers allow obtaining of the full-length  
 CC CDNA easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH1633 to AAH18742 represent human cDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.  
 XX Sequence 366 AA:  
 SO  
 Query Match 29.9%; Score 215.5; DB 22; Length 366;  
 Best Local Similarity 44.4%; Pred. No. 4.5e-18;  
 Matches 48; Conservative 17; Mismatches 40; Indels 3; Gaps 1;  
 QY 15 DEBTLWEMSHHRIVRCICPSRLIPYRQAVYLCQDDEEVLHSPRTNSAMRAGHLL 74  
 DB 6 NDECKNVVLEGFEVTLTVSDIPSRITPYRQCKVLPNDDSEQLSPDNVIRKRGVLL 65  
 QY 75 DLKTRGKNGAIAFLESIKFHNPDVYTVGLQPDVDF--NFSGES 119  
 DB 66 DILORTGHKGYVAFLESLIELYYPOLYKRYKGVKPAVFSMTIDSGES 113

RESULT 13  
 AAG76060  
 ID AAG76060 standard; protein; 366 AA.  
 AC AAG76060;  
 XX 09-MAY-2003 (first entry)  
 DE Caspase recruitment domain containing protein, CARD-10X.  
 XX CARD; caspase recruitment domain; apoptosis; cell adhesion; inflammation;  
 XX cytokine receptor signalling; cancer; glioma; carcinoma; adenocarcinoma;  
 XX CARD-containing polypeptide associated disorder; sarcoma; melanoma;  
 XX hamartoma; leukaemia; lymphoma; keratinocyte hyperplasia; neoplasia;  
 XX keloid; benign prostatic hypertrophy; inflammatory hyperplasia; fibrosis;  
 XX restenosis; allergy; arthritis; lupus; Sjogren's syndrome; sepsis;  
 XX Crohn's disease; ulcerative colitis; graft versus host disease; stroke;  
 XX abnormal cell death disease; myocardial infarction; heart failure;  
 XX neurodegenerative disease; Parkinson's disease; Alzheimer's disease; HIV;  
 XX CARD-10X; caspase activator; caspase inhibitor.  
 OS Unidentified.  
 XX Key Location/Qualifiers  
 FH 12..83  
 FT Domain /label= CARD  
 FT /note= "Caspase recruitment domain. Specifically  
 FT 154..297  
 FT Domain /label= Filament domain  
 FT /note= "Specifically claimed in claim 12"  
 XX US2002164703-A1.  
 XX 07-NOV-2002.  
 PD 19-DEC-2001; 2001US-0032159.  
 PF 21-DEC-2000; 2000US-257457P.  
 PR (PAWL/) PAWLOWSKI K.  
 PA (REED/) REED J C.  
 PA (GODZ/) GODZIK A.  
 XX Pawlowski K, Reed JC, Godzik A;  
 PI WPI; 2003-288137/28.  
 DR N-PSDB; ABX11432.  
 DR New isolated CARD-containing nucleic acids, useful for the diagnosis  
 PT and treatment of disorders with aberrant expression or activity of the  
 PT CARD-containing polypeptide, such as cancer, stroke, arthritis, heart  
 PT failure and AIDS -  
 XX Claim 13; Fig 1; 34pp; English.  
 PS The invention relates to an isolated nucleic acid molecule encoding a  
 CC caspase recruitment domain (CARD) containing polypeptide. CARD containing  
 CC polypeptides are involved in apoptosis (as caspase activators and caspase  
 CC inhibitors), cell adhesion, inflammation and cytokine receptor  
 CC signalling. The methods and compositions of the present invention are  
 CC useful for the diagnosis and treatment of disorders associated with the  
 CC aberrant expression or activity of the CARD containing polypeptide such  
 CC as cancer, glioma, carcinoma, adenocarcinoma, sarcoma, melanoma, keloid,  
 CC hamartoma, leukaemia, lymphoma, keratinocyte hyperplasia, neoplasia,  
 CC benign prostatic hypertrophy, inflammatory hyperplasia, fibrosis,  
 CC restenosis, allergy, arthritis, lupus, Sjogren's syndrome, sepsis,  
 CC Crohn's disease, ulcerative colitis, graft versus host disease, stroke,  
 CC abnormal cell death diseases such as stroke,  
 CC myocardial infarction, heart failure, neurodegenerative diseases like  
 CC Parkinson's disease and Alzheimer's disease, and HIV infection. The  
 CC present sequence represents the amino acid sequence of the caspase  
 CC recruitment domain containing protein, CARD-10X.

XX Sequence 366 AA;  
 SQ Query Match 29.9%; Score 215.5; DB 24; Length 366;  
 Best Local Similarity 44.4%; Pred. No. 4,5e-18;  
 Matches 48; Conservative 17; Mismatches 40; Indels 3; Gaps 1;

QY 15 DEETLWEMSHRRIRVRCICPSRLTPYRQAKVLCQDEEVLHSPRLTNSAMRAGHL 74  
 Db 6 NDECMNVLEGKRVLTSLVIDPSRITPYRQCKVLPDDEEVLSDPMLVIRKRVYLL 65  
 QY 75 DLIKTRKNGAIAFLBSLKFHPDVTTLVGLQPDVFS--NFGSGS 119  
 Db 66 DILORTGHKGVAFLBSLELYYPOLYKVTGKPEKPARVFSMIDAGSGS 113

RESULT 14  
 AAU01205  
 ID AAU01205 standard; Protein; 536 AA.  
 XX AAU01205;  
 AC 12-SEP-2001 (first entry)  
 XX Human caspase recruitment domain, CARD-9 polypeptide.  
 XX  
 XX Human; caspase recruitment domain; CARD-9; Bcl-10; NF-kappaB;  
 KW apoptosis; hyperproliferative disorder; autoimmune; neurological;  
 KM inflammatory disorder; viral infection; stress-related response.  
 XX Homo sapiens.  
 XX  
 FH Location/Qualifiers  
 FT 2..5  
 FT /note= "Casein kinase II phosphorylation site"  
 FT 7..98  
 FT /note= "CARD domain"  
 FT 23..26  
 FT /note= "Casein kinase II phosphorylation site"  
 FT 92..95  
 FT /note= "CAMP- and CGMP-dependent protein kinase phosphorylation site"  
 FT Modified-site  
 FT 95..97  
 FT /note= "Protein kinase C phosphorylation site"  
 FT 95..98  
 FT /note= "Casein kinase II phosphorylation site"  
 FT 138..140  
 FT /note= "Protein kinase C phosphorylation site"  
 FT 138..141  
 FT /note= "Casein kinase II phosphorylation site"  
 FT 140..416  
 FT /note= "Coiled coil domain"  
 FT 176..183  
 FT /note= "Tyrosine kinase phosphorylation site"  
 FT 197..213  
 FT /note= "Indole-3-glycerol phosphate synthase homology region"  
 FT 228..231  
 FT /note= "CAMP- and CGMP-dependent protein kinase phosphorylation site"  
 FT 231..233  
 FT /note= "Protein kinase C phosphorylation site"  
 FT 267..270  
 FT /note= "Casein kinase II phosphorylation site"  
 FT 285..338  
 FT /note= "Cysteine rich repeat homology region"  
 FT 303..305  
 FT /note= "Protein kinase C phosphorylation site"  
 FT 333..336  
 FT /note= "CAMP- and CGMP-dependent protein kinase phosphorylation site"  
 FT 363..366  
 FT /note= "Casein kinase II phosphorylation site"

FT Modified-site 425..428  
 FT /note= "Casein kinase II phosphorylation site"  
 FT Modified-site 431..433  
 FT /note= "Protein kinase C phosphorylation site"  
 FT Modified-site 450..452  
 FT /note= "Protein kinase C phosphorylation site"  
 FT Modified-site 453..458  
 FT /note= "N-myristoylation site"  
 FT Modified-site 460..462  
 FT /note= "Protein kinase C phosphorylation site"  
 FT Modified-site 481..486  
 FT /note= "N-myristoylation site"  
 FT Modified-site 483..486  
 FT /note= "Casein kinase II phosphorylation site"  
 FT Modified-site 524..527  
 FT /note= "N-glycosylation site"  
 FT Modified-site 526..529  
 FT /note= "Casein kinase II phosphorylation site"  
 FT Modified-site 527..532  
 FT /note= "N-myristoylation site"  
 FT Modified-site 531..534  
 FT /note= "Casein kinase II phosphorylation site"

WO200140468-A2.  
 PD 07-JUN-2001.  
 XX 01-DEC-2000; 2000WO-US32716.  
 PF 03-DEC-1999; 99US-0168780.  
 PR 18-FEB-2000; 2000US-0507553.  
 PR 25-FEB-2000; 2000US-0513904.  
 PR 10-OCT-2000; 2000US-0668791.  
 PA (MILL-) MILLENNIUM PHARM INC.  
 PI Bertin J;  
 XX WPI, 2001-367809/38.  
 DR N-PSDB; NMS05387.  
 XX Novel caspase recruitment domain (CARD) proteins, CARD-9, CARD-10, CARD-11, useful as targets for therapy, as immunogens, and in screening and detection assays -  
 PS Claim 9; Fig 5A-5B; 145pp; English.  
 XX The present sequence represents novel human caspase recruitment domain, CARD-9. The polynucleotide encoding this sequence was isolated from a human megakaryocyte cDNA library. Also described are novel human sequences for CARD-10 and CARD-11 (AAU01206, AAU01207) and rat CARD-9 (AAU01204). CARD-9, CARD-10 and CARD-11 interact with Bcl-10 which is thought to activate nuclear factor (NF)-kappaB and apoptosis. The sequences of the invention can be used for treating a disorder associated with abnormal levels of apoptosis by modulating the expression or activity of CARD-9, CARD-10, or CARD-11. They can be used for the treatment of hyperproliferative disorders (e.g. cancer), autoimmune disorders (e.g. systemic lupus erythematosus), neurological disorders (e.g. Alzheimer's disease), inflammatory disorders (e.g. Crohn's disease), and viral infection (e.g. HIV). The CARD polypeptide, polynucleotide and an antibody which selectively binds to CARD can be used in screening and detection assays (e.g. chromosomal mapping, tissue typing), predictive medicine (prognostic assays, monitoring clinical trials, and therapy (treatment and prophylaxis)). The CARD polypeptide may be used to screen for drugs that bind to and/or modulate it. CARD sequences are potential targets for regulating inflammation, cancer, NF-kappaB signaling, stress-related response and apoptosis in human disease. A host cell containing a polynucleotide encoding CARD can be used to create transgenic animals.

SQ Sequence 536 AA;  
 Query Match 29.9%; Score 215.5; DB 22; Length 536;



Best Local Similarity 44.4%; Pred. No. 7.9e-18;  
Matches 48; Conservative 17; Mismatches 40; Indels 3; Gaps 1;  
QY 15 DEETLWEMSHRRIRVRCICPSRLTPYLROAKVLCQDEBEVLHSPRLTNSAMRAGHLL 74  
Db 6 NDEECWNVLEGRVTLTSVIDPSRLTPYLROCKVLPDPDEQVLSDPNLVIRKRVGVL 65  
QY 75 DLKTRGKNGAIAFLSEIKFNPVYTLVTGLQPDVDF--NFGES 119  
Db 66 DILQRTGHKGYVAFLESLLELYPOLYKKVTKGKPARVFSMIIIDAGES 113  
RESULT 15  
AAG79553  
ID AAG79553 standard; Protein; 536 AA.  
XX AAG79553;  
AC AAG79553;  
XX 09-DEC-2002 (first entry)  
XX  
XX Human CARD-9.  
XX  
XX Rat; human; caspase recruitment domain; CARD-9; CARD-10;  
XX CARD-11; apoptosis; inflammation; cell growth; cell death;  
XX lymphocyte activation; cancer; melanoma; autoimmune disease;  
XX arthritis; neurological disorder; Alzheimer's disease.  
XX Homo sapiens.  
XX OS  
XX  
XX Key Location/Qualifiers  
XX Domain 7..98  
XX /label= CARD\_domain  
XX Domain 140..416  
XX /label= Coiled-coil\_domain  
XX Domain 197..213  
XX /label= Indole-3-glycerol\_phosphate\_synthase  
XX Region 285..338  
XX /label= Cysteine\_rich\_repeat\_region  
XX  
XX MO200270652-A2.  
XX  
XX PD 12-SEP-2002.  
XX  
XX PF 28-FEB-2002; 2002WO-US06147.  
XX  
XX PR 02-MAR-2001; 2001US-0798412.  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
XX  
XX Bertin J;  
XX  
XX WPI; 2002-698749/75.  
XX N-PSDB; ABA00332.  
XX  
XX CARD-9, CARD-10 or CARD-11 polypeptides and polymucleotides, useful for  
XX treating disorders associated with inappropriate apoptosis or  
XX lymphocyte activation, e.g. cancer  
XX  
XX Disclosure; Fig 5; 15pp; English.  
XX  
XX This sequence represents human caspase recruitment domain (CARD)-9.  
XX CARD proteins play roles in apoptotic and inflammatory signalling  
XX pathways. CARD-9, -10 and -11 participate in the network of  
XX interactions that modulate caspase activity. They are thought to be  
XX useful as modulating agents for regulating a variety of cellular  
XX processes including cell growth and cell death. CARD proteins and  
XX nucleic acids are useful for treating a disorder associated with  
XX inappropriate apoptosis or lymphocyte activation or for diagnosing  
XX subjects having or that are at risk of developing a disorder associated  
XX with aberrant CARD-9, CARD-10 or CARD-11 expression or activity, such  
XX as cancer e.g. melanoma, autoimmune disorders e.g. arthritis, or  
XX neurological disorders e.g. Alzheimer's disease.

SQ Sequence 536 AA;  
Query Match 29.9%; Score 215.5; DB 23; Length 536;  
Best Local Similarity 44.4%; Pred. No. 7.9e-18;  
Matches 48; Conservative 17; Mismatches 40; Indels 3; Gaps 1;  
QY 15 DEETLWEMSHRRIRVRCICPSRLTPYLROAKVLCQDEBEVLHSPRLTNSAMRAGHLL 74  
Db 6 NDEECWNVLEGRVTLTSVIDPSRLTPYLROCKVLPDPDEQVLSDPNLVIRKRVGVL 65  
QY 75 DLKTRGKNGAIAFLSEIKFNPVYTLVTGLQPDVDF--NFGES 119  
Db 66 DILQRTGHKGYVAFLESLLELYPOLYKKVTKGKPARVFSMIIIDAGES 113  
Search completed: February 18, 2004, 04:44:28  
Job time : 77 secs

OM protein - protein search, using SW model

Run on: February 18, 2004, 04:37:37 ; Search time 36 Seconds  
(without alignments)  
371.318 Million cell updates/sec

Title: US-10-032-159A-16  
Perfect score: 720  
Sequence: 1 MGEICRDSALALDEETLW.....SPEDGIAGTSRNLRLVTPX 139

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : PIR\_76:\*

```
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	81	11.2	579	2	F69157	exonuclease ABC c
2	80	11.1	2114	2	E59505	hypothetical protein
3	76.5	10.6	532	2	H83493	hypothetical protein
4	75	10.4	271	1	A89124	probable phosphoenolpyruvate carboxylase
5	74	10.3	210	2	S55671	hypothetical protein
6	73.5	10.2	315	2	T26673	hypothetical protein
7	73.5	10.2	1113	2	T14260	period protein Per1
8	73	10.1	265	2	H97280	hydroxyethylthiazole synthase
9	72.5	10.1	1318	2	T21265	hypothetical protein
10	72.5	10.1	1583	2	S59644	sister chromatid cohesion factor
11	72	10.0	299	2	I37468	housekeeping gene 1
12	72	10.0	1234	2	B36186	factor protein 2
13	71.5	9.9	660	2	T45559	receptor protein R
14	71	9.9	213	2	A90269	conserved hypothetical protein
15	71	9.9	265	2	T33012	hypothetical protein
16	71	9.9	428	2	AP0241	probable coenzyme A synthase
17	70.5	9.8	263	2	A30227	hypothetical protein
18	70.5	9.8	329	2	H83604	hypothetical protein
19	70.5	9.8	356	2	S13221	GTP-binding regulatory protein
20	70.5	9.8	513	2	S69181	protein disulfide isomerase
21	70.5	9.8	2376	2	S48405	probable membrane protein
22	70	9.7	599	2	AB3375	beta-(1->2)glucanase
23	69.5	9.7	349	2	AF2110	iron(III) dicarboxylate reductase
24	69.5	9.7	522	2	D72349	conserved hypothetical protein
25	69.5	9.7	1452	1	S17670	protein-tyrosine phosphatase
26	69	9.6	496	2	E83849	spore germination protein
27	69	9.6	537	2	A30325	membrane alanine aminotransferase
28	68.5	9.5	532	2	A65138	hypothetical protein
29	68.5	9.5	532	2	A91162	probable 2-component system

30	68.5	9.5	532	2	B8608	probable 2-component
31	68.5	9.5	557	2	A12239	hypothetical protein
32	68.5	9.5	1115	2	T13955	period protein Per
33	68	9.4	603	2	S34130	serine/threonine-S
34	67.5	9.4	270	2	T50952	transposase relat
35	67.5	9.4	460	2	E86922	probable FAD-linker
36	67.5	9.4	521	2	AB0269	anthranilate synth
37	67.5	9.4	872	2	E71852	valine-tRNA ligase
38	67.5	9.4	882	1	IJHUC2	cadherin 1 precursor
39	67.5	9.4	1470	2	B57062	SMB9 protein - year
40	67	9.3	404	2	F81724	conserved hypothec
41	67	9.3	493	2	T04961	hypothetical prote
42	67	9.3	728	2	S57142	hypothetical prote
43	67	9.3	867	2	H70411	alanine-tRNA ligas
44	67	9.3	1012	2	E90389	conserved hypothet
45	67	9.3	1446	2	T11018	hypothetical prote

## ALIGNMENTS

```

RESULT 1
F69157
excinnuclease ABC chain C - Methanobacterium thermoautotrophicum (strain Delta H)
C/Species: Methanobacterium thermoautotrophicum
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Apr-2000
C/Accession: F69157
R/Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Qiu, D.; Spadatorra, R.; Vicaire, R.; Wang, Y.; Wierbowski, J.; Gibson, R.; Ujwand, N.
K1, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A/Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functio
A/Reference number: A69000; M01D:98037514; PMID:9371463
A/Accession: F69157
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-579 <MTH>
A/Cross-references: GB:AE000828; GB:AE000666; NID:g2621504; PIDN:AA84947.1; PID:g2621505C
A/Experimental source: strain Delta H
C/Genetics:
A/Gene: MTH441
A/Start codon: TTG
C/Superfamily: excinnuclease ABC chain C

Query Match      11.2%; Score 81; DB 2; Length 579;
Best Local Similarity 27.6%; Pred. No. 3.3;
Matches 43; Conservative 26; Mismatches 53; Indels 34; Gaps 9;

```

```

Query Match      11.2%; Score 81; DB 2; Length 579;
Best Local Similarity 27.6%; Pred. No. 3.3;
Matches 43; Conservative 26; Mismatches 53; Indels 34; Gaps 9;

QY          7 RDSALTAIDRETL-----MEMMES---HHHRTVRCICPSRLTPYLARQ-----AKVLGOLD 53
           || :|||:||:||||:| | | | | | | | | | | | | | | | | | | | | |
Db          262 RDGKITGKDPIFLNGSAPRTEILLEAFLLKQYYA-PRRPSSILTYQPEVDGVIAWLSLR 321

QY          54 EEEV-LHSP-----RLTNSAMRAGHLDLTKRGKGALAFPLESLFHNPDVYLTITGL 106
           ||| :|||:| | | | | | | | | | | | | | | | | | | | | |
Db          322 GEEVKHSPPGGAGRRRLNTAWTASVYLKKQKGVADALQLKDDKL--PETIRRMGL 379

QY          107 QPDVDFSNFGSESSD-----FDG--LAGTSRNRL 134
           || | | | | | | | | | | | | | | | | | | | | | |
Db          380 ----DISNIGESATGSVAVFIDGKPSSGSGSYRRRI 411

RESULT 2
E96505
hypothetical protein T7023.25 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revansion 02-Mar-2001 #text_change 31-Mar-2001
C/Accession: E96505
R/Theologis, A.; Becker, J.R.; Palm, C.U.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hultzer, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
A.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luross, J.S.; Maitl, R.; Marzialis,

```

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzman, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 Ker, M.; Wu, D.; Xu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: E96505  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-2114 <STO>  
 A:Cross-references: GB:AE005173; NID:g11120821; PIDN:AA031000.1; GSPDB:GN00141  
 C:Genetics:  
 A:Gene: T7023.25  
 A:Map position: 1

Query Match 11.1%; Score 80; DB 2; Length 2114;  
 Best Local Similarity 26.5%; Pred. No. 21;  
 Matches 40; Conservative 22; Mismatches 69; Indels 20; Gaps 5;  
 Db 4 LCRDSDALNALDE-ETLWEMESHHR-----YRCICPRLTPYLRQAATVQL 52  
 1206 LSPDSTETVSELSLSPSPITRHKTAISSKQIGIHLASRTY-NAATVCEL 1264  
 Qy DEEEVHSPRLTNSAMRAGHLLDLKTRGKGAIAFLESK-----FENPDVTLVTGL 106  
 Db 1265 FSEEHRLDELANKAL--SPLEMTNTLESERVALTLVLTWGINRPDIITSLGN 1322  
 Qy 107 QPDVDFSNFSGSSDPDLGAGTSRNLRLVT 137  
 Db 1323 PLDNIYKILSDSSLESKTSARICRPLFT 1353

RESULT 3  
 H83493  
 hypochetrical protein PA1214 [imported] - *Pseudomonas aeruginosa* (strain PA01)  
 C:Species: *Pseudomonas aeruginosa*  
 C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C:Accession: H83493  
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kae, A.; Lardig, K.; Lim,  
 .; Lory, S.; Olson, M.V.  
 A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho  
 A:Reference number: A82950; MUID:20437337; PMID:10984043  
 A:Accession: H83493  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-532 <STO>  
 A:Cross-references: GB:AE004551; GB:AE004091; NID:g9947135; PIDN:AA04603.1; GSPDB:GN001  
 A:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: PA1214

Query Match 10.6%; Score 76.5; DB 2; Length 532;  
 Best Local Similarity 26.7%; Pred. No. 8.5;  
 Matches 40; Conservative 18; Mismatches 67; Indels 25; Gaps 6;  
 Qy 4 LCRDSDALNALDE-ETLWEMESHHRIVTC-----ICPRLTPYLRQA 45  
 Db 197 LCRDSDALNALDE-ETLWEMESHHRIVTC-----ICPRLTPYLRQA 255  
 Qy 46 AKYLCOLDEEVHSPRLTNSAMRAGHLLDLKTRGKGAIAFLESKFNHPDVTLVT 104  
 Db 256 FHLCEGDEESLPHRLQFELRQAFMLPILRAVNGATRMSSLLMYG----RLAD 311  
 Qy 105 GLQPDVDFSNFSGSSDPDLGAGTSRNLRL 134  
 Db 312 GIGGGYHCVTLGEGAD-ELFMGYPRHLEL 340

RESULT 4  
 A89124  
 probable phosphoesterase (BC 3.1.-.-) K07C11.7 [similarity] - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*

C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 02-Nov-2001  
 C:Accession: A89124  
 R:Anonymous. The C. elegans Sequencing Consortium.  
 A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biolo  
 A:Reference number: A75000; MUID:99069613; PMID:9851916  
 A:Note: see websites genome.muc1.edu/sec/C.elegans/ and www.sanger.ac.uk/Projects/C.el  
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; am  
 A:Accession: A89124  
 A:Molecule type: DNA  
 A:Residues: 1-271 <STO>  
 A:Cross-references: GB:chr V; PIDN:AA96171.1; PID:g1255827; GSPDB:GN00023; CESP:K07C11  
 C:Comment: This sequence has motifs characteristic of a variety of phosphoesterases.  
 C:Genetics:  
 A:Gene: K07C11.7  
 A:Map position: 5  
 C:Superfamily: *Caenorhabditis elegans* probable phosphoesterase C25R10.12; phosphoesterase  
 C:Keywords: hydrolase  
 F:68-128/Domains: phosphoesterase core homology <PFC>

Query Match 10.4%; Score 75; DB 1; Length 271;  
 Best Local Similarity 26.5%; Pred. No. 5.2;  
 Matches 36; Conservative 19; Mismatches 61; Indels 20; Gaps 6;  
 Qy 14 LDEETLWEMESHHRIVRCIC-----PSRLTPYLRQAATVQLDEEVHSPR 62  
 Db 27 IDPDAENELWDSIHTYQNIIVERKALGINSPTDTPYL--KVCISDTHQLNVT 83  
 Qy 63 LTNSAMRAGHLLDLKTRGKGAIAFLESKFNHPDVTLVTGLQPDVDFSN--FSGES 119  
 Db 84 VPDGDVLI-HAGDTNNKKEELIKENETRF--PKYKLVAGNHELGFDDENQER 140  
 Qy 120 SDFDLGAGTSRNLRL 135  
 Db 141 QDADKXGIGTEGVYIL 156

RESULT 5  
 S55671  
 hypochetrical protein E10 - equine herpesvirus 2  
 C:Species: equine herpesvirus 2  
 C>Date: 27-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 24-Nov-1995  
 C:Accession: S55671  
 R:Teiford, E.A.R.; Watson, M.S.; Aird, H.C.; Perry, J.; Davison, A.J.  
 J. Mol. Biol. 249, 520-528, 1995  
 A:Title: The DNA sequence of equine herpesvirus 2.  
 A:Reference number: S55594; MUID:95302501; PMID:7783207  
 A:Accession: S55671  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-210 <TEL>  
 A:Cross-references: GB:U20824; NID:g695172; PIDN:AA013865.1; PID:g695250  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1995  
 C:Superfamily: equine herpesvirus 2 hypochetrical protein E10

Query Match 10.3%; Score 74; DB 2; Length 210;  
 Best Local Similarity 28.0%; Pred. No. 4.7;  
 Matches 30; Conservative 14; Mismatches 27; Indels 36; Gaps 6;  
 Qy 2 GELCRDSDALNALDEETLWEM-----MESHRRIVRCICPSRLTPYLRQA 46  
 Db 10 GDPC-----VTLREEDIMVERLCLEKRLVLYSHKSHK-----LDHRAK 52  
 Qy 47 KYLCOLDEEVHSPRLTNSAMRAGHLLDLKTRGKGAIAFLESK 93  
 Db 53 KILSRDAEY--SSRAI--GRSRAGLIVDMCQDHR--GFOLKESCK 95

RESULT 6  
 T6673  
 hypochetrical protein Y3810A.w - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 20-Jun-2000

C:\Accession: T26673  
R:\Wallis, J.  
submitted to the EMBL Data Library, September 1999  
A:\Reference number: Z20252  
A:\Accession: T26673  
A:\Status: preliminary; translated from GB/EMBL/DDBJ  
A:\Molecule type: DNA  
A:\Residues: 1-315 <WT>  
A:\Cross-references: EMBL:AL110484; PIDN:CAB5444.1; CESP:X38E10A.w  
A:\Experimental source: clone Y38E10A  
C:\Genetics:  
A:\Gene: CESP:X38E10A.w  
A:\Introns: 112/1; 172/2  
C:\Superfamily: Caenorhabditis elegans hypothetical protein Y38E10A.w

Query Match            10.2%; Score 73.5; DB 2; Length 315;  
Best Local Similarity 27.0%; Pred.No. 8.9;  
Matches 31; Conservative 15; Mismatches 42; Indels 27; Gaps 5;

Dy      23 MESHRHIVRCICPSRL-TPLYROAKVLCQDEBEVLHSPRLTNSAMRA-----GHLLD 75  
       :  
Db      18 LAENRPPEQQLADSLPFTAYMKLKQAEL-----SPLYSALISALSISRHLPA 70  
       :  
Dy      76 LLKTRGNKGALA---FLESLEFNPDV-----YTLVTGHPDPVDFFSNFG 117  
       :  
Db      71 YLRQANKNGLATATEKLIOSTKDHNAAAGMTATFLITLVKKNPVPFPADLOG 125  
       :

RESULT 7  
T14260  
period protein Per3 - mouse  
C:\Species: Mus musculus (house mouse)  
C:\Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
C:\Accession: T14260  
R:\Zylka, M.J.; Shearman, L.P.; Weaver, D.R.; Regpert, S.M.  
Neuron 20, 1103-1110, 1998  
A:\Title: Three period homologs in mammals: differential light responses in the suprachiasmatic nucleus  
A:\Reference number: Z17943; MWID:96318231; PMID:9655499  
A:\Accession: T14260  
A:\Status: preliminary; translated from GB/EMBL/DDBJ  
A:\Residues: 1-1113 <ZYL>  
A:\Molecule type: mRNA  
A:\Cross-references: EMBL:AF050182; NID:g3136149; PID:g3136150; PIDN:AAC40147.1  
C:\Genetics:  
A:\Gene: Per3  
C:\Keywords: circadian rhythm

Query Match            10.2%; Score 73.5; DB 2; Length 1113;  
Best Local Similarity 25.2%; Pred.No. 44;  
Matches 26; Conservative 19; Mismatches 33; Indels 25; Gaps 4;

Dy      6 RRDNALDLDBETTWEMESHRRIRVICIPSRILTYIROAKVLCQ-----LDIEEVLT 58  
       :  
Db      1003 QRDALPGAAEESSIMWME----RTPECVMITYGVPRGRREVLYKODLEKLSWEGCOPL 1058  
       :  
Dy      59 HSPRL-----TNSMRAGHLLDLL-KTRGNKAIA 87  
       :  
Db      1059 FSPQRREELAENVRSWHSHTAPQEGHQSCVACEBDRGSVGDTA 1101  
       :

RESULT 8  
H97280  
hydroxyethylcholine kinase Thik/Thik (fst) CAC3095 [Imported] - Clostridium acetobutylicum  
C:\Species: Clostridium acetobutylicum  
C:\Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 30-Sep-2001  
R:\Noelling, J.; Berton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
J.; Daly, M.J.; Bennett, G.N.; Koehn, E.V.; Smith, D.R.  
A:\Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld  
A:\Reference number: A96900; MWID:21359325; PMID:21359325  
A:\Accession: H97280  
A:\Status: preliminary

```

A:Molecule type: DNA
A:Residues: 1-265 <KIR>
A:Cross-references: GB:AE001437; PIDN:AAK81035.1; PID:g15026160; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC3095
C:Superfamily: phosphomethylpyrimidine phosphate kinase

Query Match      10.1%; Score 73; DB 2; Length 265;
Best Local Similarity 23.8%; Pred. No. 8.1;
Matches 20; Conservative 23; Mismatches 29; Indels 12; Gaps 4;

Oy 29 RIVRCICP--SRITPYRQAKVLCQL---DEEVLHSPRTNS-----AMRAGHLID-L 76
Db 125 KLTICLPLPLLTITPNIPEAAVLCGFKIKSEQDWMBAKTSKINVGITIKGHLVDNA 184
Oy 77 LKTRGNKAIPAFLESLKFNPDVY 100
Db 185 TDLFYKNGYIEWFKLEKIDNPNT 208

RESULT 9
T21266
hypochemical protein P22E10.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T21266
R:Gardner, A.
submitted to the EMBL Data Library, November 1995
A:Reference number: Z19398
A:Accession: T21266
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1318 <WIL>
A:Cross-references: EMBL:Z67882; PIDN:CAA31799.1; GSPDB:GN00028; CESP:F22E10.1
A:Experimental source: clone P22E10
C:Genetics:
A:Gene: CESP:F22E10.1
A:Map position: X
A:Insertions: 26/3; 81/3; 114/2; 156/3; 201/2; 258/3; 300/2; 396/3; 676/2; 818/3; 939/3; 10
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match      10.1%; Score 72.5; DB 2; Length 1318;
Best Local Similarity 26.3%; Pred. No. 69;
Matches 25; Conservative 13; Mismatches 32; Indels 23; Gaps 4;

Oy 57 VLHS----PLTNSAMRAGHLIDLKTRGNGLI-----AFLESLKFNPDVYT 101
Db 1032 VMSSSHNPFEPFAKATPAGMLFRLITRKSTGDMEGNNTETRGNVLPFSVKFSYFQ--- 1088
Oy 102 LVTLGLCP---DVDFSNFGSGSSDPDGLAGTSRN 131
Db 1089 --RMPQVMTDLHFSAHSGQTVALVGPSSGSGKS 1119

RESULT 10
S59644
sister chromatid cohesion molecule Wis4p - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 14-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 18-Feb-2000
C:Accession: T38603; T43392; S59644
R:Devlin, K.; Churche, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, July 1995
A:Reference number: Z21731
A:Accession: T38603
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1583 <DE2>
A:Cross-references: EMBL:Z50113; NID:g914878; PIDN:CAA90463.1; PID:g914883; GSPDB:GN0006
A:Experimental source: strain 972h-; cosmid c31A2
R:Puriya, K.; Takahashi, K.; Yanagida, M.
submitted to the EMBL Data Library, August 1998
A:Description: Faithful anaphase is ensured by Wis4, a sister chromatid cohesion molecule

```

A:Reference number: Z22478  
 A:Accession: T43392  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1583 <FUR>  
 A:Cross-references: EMBL:AB016866; PIDN:CAB19489.1  
 C:Genetics:  
 A:Gene: mls4, SPAC31A2.05C  
 A:Map position: 1  
 A:introns: 33/1, 98/2, 543/3, 699/3, 1294/2, 1339/3, 1558/3

Query Match 10.1%; Score 72.5; DB 2; Length 1583;  
 Best Local Similarity 29.0%; Pred. No. 87;  
 Matches 31; Conservative 14; Mismatches 39; Indels 23; Gaps 6;

QY 11 LTALDEETLWEMSHRRIRVRCIP--SRITPYLRQAKVL--CQLDEEYVLSHPRLTNS 66  
 DB 1091 LTKAGTATLME-----IVPCLCSLFTRLNDYERLKKIVSCSKSIEARHS---ENN 1139

QY 67 AMRAGHLIDLTKRGKNGAI-----AFLESLEKFNPP--DYVTLVTVG 105  
 DB 1140 FQKWRLLIDLIGFRYGDLMRINDWKHSUDFISPCDDAYVILG 1186

RESULT 11  
 137468  
 house keeping gene 33 - human  
 C:Species: Homo sapiens (man)  
 C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 18-Aug-2000  
 C:Accession: 137468

R.Braun, A.; Kammerer, S.; Weissenhorn, W.; Weiss, E.H.; Cleve, H.  
 Gene 146, 291-295, 1994  
 A:Title: Sequence of a putative human housekeeping gene (HK33) localized on chromosome 1  
 A:Reference number: 137468; M01D:94357452; PMID:8076834  
 A:Accession: 137468

A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-239 <RES>  
 A:Cross-references: EMBL:X75535; NID:g1644300; PIDN:CA53225.1; PID:g551250  
 C:Genetics:  
 A:Gene: hk 33  
 C:Superfamily: Caenorhabditis elegans F54F2.8 protein

Query Match 10.0%; Score 72; DB 2; Length 299;  
 Best Local Similarity 26.0%; Pred. No. 12;  
 Matches 32; Conservative 20; Mismatches 51; Indels 20; Gaps 6;

QY 1 MDELGRDSALTADE--ETLWEMSHRRIRVRCIPSRITPYLRQAKVLCQLDEEYV 58  
 DB 179 MONTLSKDVLYPSLKEITKTPSWTQSHRESLP---PQGFKYQEQSHSWCKICEQEA 234

QY 59 HSPRLTNSAMRAGH--LTDLTKRGKNGAIAFLSKFHPNDVYV--LVYGLQPDVDFSN 114  
 DB 235 EPTDSETQKARFENVDLMQ-----LQDLG-HPPKELAGMPPGLNPLDLALN 284

QY 115 FSG 117  
 DB 285 LSG 287

RESULT 12  
 B36186  
 I factor protein 2 - fruit fly (Drosophila teissleri) transposon I  
 C:Species: Drosophila teissleri  
 C:Date: 14-Dec-1990 #sequence\_revision 14-Dec-1990 #text\_change 31-Mar-2000  
 C:Accession: B36186

R.Abad, P.; Vaurio, C.; Pelisson, A.; Chaboisier, M.C.; Buseau, I.; Bucheton, A.  
 Proc. Natl. Acad. Sci. U.S.A. 86, 8887-8891, 1989  
 A:Title: A long interspersed repetitive element--the I factor of Drosophila teissleri--  
 A:Reference number: A36186; M01D:90046894; PMID:2554335  
 A:Accession: B36186

A:Status: preliminary  
 A:Molecule type: DNA

A:Residues: 1-1234 <ABA>  
 A:Cross-references: GB:W28878; NID:g157750; PID:g950316  
 C:Genetics:  
 A:Gene: FlyBase:Dmel/I-element  
 A:Cross-references: FlyBase:Fgn0013017

Query Match 10.0%; Score 72; DB 2; Length 1234;  
 Best Local Similarity 27.1%; Pred. No. 71;  
 Matches 26; Conservative 20; Mismatches 34; Indels 16; Gaps 4;

QY 59 HSPRLTNSAMRAGHIDLTKT--RGKNG-----AIAFLSKL-----FHPNDVYVTVGLQ 107  
 DB 979 YSSVLTSETTALIEALELITKTRGKGIWFDSDAISIDSKPNNSFPYPRILSLITQLA 1038

QY 108 PDVDFSNFSGESDF-----DGLAGTSRLRLVTP 138  
 DB 1039 FKIKIMTIPGHSIGTGNELADQAKLASNPILVTP 1074

RESULT 13  
 T45569  
 receptor protein kinase-like protein - Arabidopsis thaliana  
 N:Alternate names: protein F11C1.70  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Mar-2000  
 C:Accession: T45569

R.Bargues, M.; Collado, M.C.; Navarro, P.; Terol, J.; Perez-Alonso, M.; Mewes, H.W.; May  
 submitted to the Protein Sequence Database, December 1999  
 A:Reference number: Z23007  
 A:Accession: T45569

A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-660 <BAR>  
 A:Cross-references: EMBL:AI132976  
 A:Experimental source: cultivar Columbia; BAC clone F11C1  
 C:Genetics:  
 A:Map position: 3  
 A:introns: 472/1  
 A:Note: F11C1.70  
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol.

Query Match 9.9%; Score 71.5; DB 2; Length 660;  
 Best Local Similarity 27.0%; Pred. No. 36;  
 Matches 38; Conservative 21; Mismatches 41; Indels 41; Gaps 8;

QY 20 WEMSHRRIRVRCIP-----SRITPYLRQAKVLCQLDEEYVLSHPRLTNSAMRAGHL 74  
 DB 66 WRGVDGQDVRVRLILDGVLGSGFSF-----ETLSRLDQLRVL---SLNNST-SGSIP 116

QY 75 DL-----LK--TRKNGAIAFLSKFHPNDVYVTVGLQPDVDFSNFSGE----- 118  
 DB 117 DLSPLVNLKTLTKSKNFGSLSS-----SILSRRLTELDSFNNFSGEIPSGINALS 170

QY 119 -----SSDFDGLAGTSRLRL 134  
 DB 171 RLSSLNLEFNRNLNGTLPPLNL 191

RESULT 14  
 A90269  
 conserved hypothetical protein (imported) - Sulfolobus solfataricus  
 C:Species: Sulfolobus solfataricus  
 C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001  
 C:Accession: A90269

R.She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aways, M.U.; Chan  
 Jang, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder,  
 J.; Riet, R.A.; Ragan, M.A.; Semsen, C.W.; Van der Oost, J.  
 Submitted to GenBank, April 2001  
 A:Description: Sulfolobus solfataricus complete genome.  
 A:Reference number: A90269  
 A:Accession: A90269

A:Status: preliminary  
 A:Molecule type: DNA



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 18, 2004, 03:29:02, Search time 24 Seconds

(without alignments)  
272.363 Million cell updates/sec

Title: US-10-032-159A-16

Perfect score: 720  
Sequence: 1 MSELCRDSALTRALDEETLM.....SDFDLGAGTSNRLRLVTPX 139

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	617.5	85.8	1004	1	CARE_HUMAN
2	516.5	71.7	999	1	CARE_MOUSE
3	270.5	37.6	1147	1	CARB_HUMAN
4	252	35.0	1021	1	CARA_MOUSE
5	246	34.2	1032	1	CARA_HUMAN
6	223.5	31.0	536	1	CAR_MOUSE
7	215.5	29.9	536	1	CAR_HUMAN
8	83	11.5	233	1	BCLA_HUMAN
9	81	11.2	233	1	BCLA_PAT
10	81	11.2	579	1	UVRC_METH
11	80	11.1	965	1	AMPN_RABIT
12	79	11.0	519	1	KCG4_HUMAN
13	78	10.8	233	1	BCLA_MOUSE
14	75.5	10.5	685	1	STMI_HUMAN
15	75	10.4	220	1	YMSO_CAEEL
16	75	10.4	467	1	TMII_MOUSE
17	73.5	10.2	1113	1	PER3_MOUSE
18	72.5	10.1	1583	1	MIS4_SCHPO
19	72	10.0	299	1	PIF_HUMAN
20	71	9.9	379	1	Y452_METH
21	70.5	9.8	356	1	GBA2_CAEEL
22	70.5	9.8	513	1	TAO3_YEAST
23	70.5	9.8	2376	1	PDI_MAIZE
24	69.5	9.7	1452	1	PTPM_MOUSE
25	69	9.6	965	1	AMPN_MOUSE
26	69	9.6	966	1	AMPN_HUMAN
27	68.5	9.5	532	1	RTRC_ECOLI
28	68.5	9.5	468	1	AMPN_FELCA
29	68	9.4	468	1	TMII_HUMAN
30	68	9.4	603	1	PLK1_HUMAN
31	67.5	9.4	872	1	SYV_HELPJ
32	67.5	9.4	882	1	CAD1_HUMAN
33	67.5	9.4	1420	1	SRB9_YEAST

34	67	9.3	231	1	PYRF_OCEIH
35	67	9.3	728	1	YV89_YEAST
36	67	9.3	867	1	SYA_AQUY
37	67	9.3	902	1	GCP2_HUMAN
38	66.5	9.2	841	1	PSPI_YEAST
39	66.5	9.2	2245	1	MYSJ_DICDI
40	66	9.2	326	1	PDLI_MOUSE
41	66	9.2	524	1	HY11_BRAJA
42	66	9.2	578	1	VO22_FOMPV
43	66	9.2	853	1	NUG2_RHIME
44	66	9.2	871	1	SYA_AQUY
45	66	9.2	903	1	GCP2_MOUSE

## ALIGNMENTS

RESULT 1  
ID CARE\_HUMAN STANDARD; PRT, 1004 AA.  
AC 09XKL6; 09HVB5;  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Caspase recruitment domain protein 14 (CARD-containing MAGUK protein 2) (Carna 2).  
GN CARD14 OR CARMA2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21192234; PubMed=11278692;  
RA Berlin J., Wang L., Guo Y., Jacobson M.D., Poyet J.-L.,  
RA Srinivasula S.M., Merriam S., Disfieno P.S., Alnemri E.S.;  
RT "CARD11 and CARD14 are novel caspase recruitment domain  
RT (CARD)/membrane-associated guanylate kinase (MAGUK) family members  
RT that interact with Bcl10 and activate NF-kappaB.";  
RL J. Biol. Chem. 276:11877-11882(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21255663; PubMed=11356195;  
RA Gaide O., Martignon F., Michau O., Bonnet D., Thome M., Tschopp J.;  
RT "Carnal, a CARD-containing binding partner of Bcl10, induces Bcl10  
RT phosphorylation and NF-kappaB activation.";  
RL FEBS Lett. 496:121-127(2001).  
RN [3]  
RP ERRATUM.  
RA Gaide O., Martignon F., Michau O., Bonnet D., Thome M., Tschopp J.;  
RL FEBS Lett. 505:198-198(2001).  
RN [4]  
RP SEQUENCE OF 1-740 FROM N.A.  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strusberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.B., Scheetz T.E.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hilyk S.W.,  
RA Villalón D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield V.S.N., Krzywinski M.T., Skalska U., Smalilus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length



```

RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -1- FUNCTION: Activates NF-kappaB via Bcl10 and IKK. Stimulates the
CC phosphorylation of Bcl10.
CC -1- SUBUNIT: CARD14 and Bcl10 bind to each other by CARD-CARD
CC interaction.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: Expressed in placenta. Also detected in HeLa
CC S3 cells, but not in the other cancer cell lines tested.
CC -1- SIMILARITY: Contains 1 CARD domain.
CC -1- SIMILARITY: Contains 1 PDZ/DHR domain.
CC -1- SIMILARITY: Contains 1 guanylate kinase-like domain.
CC -1- CAUTION: Supposed to contain a SH3 domain which is not detected by
CC PROSITE, Pfam or SMART.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, AF322642; AAG5403.1; -.
CC EMBL, AY032927; AAK54453.1; -.
CC EMBL, BC018142; AAH18142.1; -.
CC EMBL, BC001326; AAH01326.1; ALT_INIT.
CC Genew: HGNC:16446; CARD14.
CC MIM: 607211; -.
CC GO: GO:0005886; C:plasma membrane; NAS.
CC GO: GO:0005515; F:protein binding activity; NAS.
CC GO: GO:0007250; P:activation of NF-kappaB-inducing kinase; NAS.
CC InterPro: IPR001315; CARD.
CC InterPro: IPR000619; Guanylate_kin.
CC InterPro: IPR001478; PDZ.
CC SMART; SM00282; Gukc; 1.
CC SMART; SM00282; PDZ; 1.
CC PROSITE; PS50209; CARD; 1.
CC PROSITE; PS500856; GUANYLATE KINASE_1; FALSE_NEG.
CC PROSITE; PS50052; GUANYLATE_KINASE_2; 1.
CC PROSITE; PS50106; PDZ; 1.
CC Coiled coil.
CC FT DOMAIN 15 107 CARD.
CC FT DOMAIN 128 409 COILED COIL (POTENTIAL).
CC FT DOMAIN 568 658 PDZ.
CC FT DOMAIN 858 990 GUANYLATE KINASE.
CC FT DOMAIN 619 671 DYKASEPLFKATVEDTTEEAVALRRVDFGCSVKYNTD
CC GYKRLDLEAK -> SRAPLSLGILMGTVAAAGVQAD
CC FTSPRCSTLTGMASALSMADVKRAHL (IN REF. 2;
CC AAH01326).
CC SQ SEQUENCE 1004 AA; 113299 MW; 74698856BE06073 CRC64;
CC -----
CC Query Match 85.8%; Score 617.5; DB 1; Length 1004;
CC Best Local Similarity 90.4%; Pred. No. 1.1e-53;
CC Matches 123; Conservative 4; Mismatches 6; Indels 3; Gaps 2;

```

```

DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Caspase recruitment domain protein 14 (Bcl10-interacting MAGUK protein
DE 2) (Bimp2).
GN CARD14 OR BIMP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21391892; PubMed=1187339;
RA McAllister-Lucas L.M., Inohara N., Lucas P.C., Ruland J., Benito A.,
RA Li O., Chen S., Chen F.F., Yamaoka S., Verma I.M., Mak T.W.,
RA Nunez G.;
RA "Bimp1, a MAGUK family member linking protein kinase C activation to
RA Bcl10-mediated NF-kappa B induction."
RL J. Biol. Chem. 276:30589-30597(2001).
RN [2]
RP SEQUENCE OF 82-743 FROM N.A.
RC TISSUE=Breast;
RX MEDLINE=22388257; PubMed=12477992;
RA Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marinska K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mollath S.J.,
RA Bosak S.A., McKernan K.J., Malek J.A., Guarnette P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skalek U., Smalins D.E.,
RA Schnerch A., Schein J.E., Jones S.J.W., Maria W.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -1- FUNCTION: Activates NF-kappaB via Bcl10 and IKK.
CC -1- SUBUNIT: CARD14 and Bcl10 bind to each other by CARD-CARD
CC interaction (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Contains 1 CARD domain.
CC -1- SIMILARITY: Contains 1 PDZ/DHR domain.
CC -1- SIMILARITY: Contains 1 guanylate kinase-like domain.
CC -1- CAUTION: Supposed to contain a SH3 domain which is not detected by
CC PROSITE, Pfam or SMART.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, AF363457; AAK60137.1; -.
CC EMBL, BC004692; AAH04692.1; -.
CC MGD; MGI:2386258; Card14.
CC InterPro: IPR001315; CARD.
CC InterPro: IPR000619; Guanylate_kin.
CC InterPro: IPR001478; PDZ.
CC Pfam; PF00595; PDZ; 1.
CC SMART; SM00282; PDZ; 1.
CC PROSITE; PS50209; CARD; 1.
CC PROSITE; PS500856; GUANYLATE KINASE_1; FALSE_NEG.
CC PROSITE; PS50052; GUANYLATE_KINASE_2; 1.
CC PROSITE; PS50106; PDZ; 1.

```

KW Coiled coil. 15 CARD.  
 FT DOMAIN 125 411 COILED COIL (POTENTIAL).  
 FT DOMAIN 572 655 PDZ.  
 FT DOMAIN 854 986 GUANYLATE KINASE.  
 FT CONFLICT 736 743 OAOQOLIA -> HLLDDHRS (IN REF. 2).  
 SQ SEQUENCE 999 AA; 113496 MW; D18350DA12430255 CRC64;

Query Match 71.7%; Score 516.5; DB 1; Length 999;  
 Best Local Similarity 77.2%; Pred. No. 1,3e-43;  
 Matches 105; Conservative 9; Mismatches 19; Indels 3; Gaps 2;

QY 1 MGELCRDSALFALDEBTLMEMESHRIRVICRCPRLTPYLRQAKVLCQDEEVLHS 60  
 DB 1 MALCLMDSTLTLDDEMLMDLHSHRCRIYOSICSRILTPYLRQAKVLCQDEEVLHS 60

QY 61 PRITNSAMRAGHLLDLKTRGNKGAIAFESLKFHNPVYTLVTGLQPDVDFSNFSG--E 118  
 DB 61 SRFETNSAMRAGHLLDLKTRGNKGAIAFESLKFHNPVYTLVTGLQPDVDFSNFSG 120

QY 119 SSDF-DGLACTSKNR 133  
 DB 121 TSKLRECLAGALISLQ 136

RESULT 3  
 CARB\_HUMAN STANDARD; PRT; 1147 AA.  
 ID CARB\_HUMAN STANDARD; PRT; 1147 AA.  
 AC Q9BXL7;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Caspase recruitment domain protein 11 (CARD-containing MAGUK protein 3) (Carna 1).  
 GN CARD11 OR CARML.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21192234; PubMed=11278692;  
 RA Berlin U., Wang L., Guo Y., Jacobson M.D., Poyet J.-L.,  
 RA Srinivasula S.M., Merriam S., Distefano P.S., Alnemri E.S.,  
 RT "CARD11 and CARD14 are novel caspase recruitment domain (CARD)/membrane-associated guanylate kinase (MAGUK) family members that interact with Bcl10 and activate NF-kappaB.";  
 RT J. Biol. Chem. 276:11877-11882(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21255663; PubMed=11356195;  
 RA Gaide O., Martinon F., Michaux O., Bonnet D., Thome M., Tschopp J.,  
 RA "Carnal, a CARD-containing binding partner of Bcl10, induces Bcl10 phosphorylation and NF-kappaB activation.";  
 RT Phosphorylation of Bcl10.  
 RT FEBS Lett. 496:121-127(2001).  
 RN [3]  
 RP ERRATUM.  
 RA Gaide O., Martinon F., Michaux O., Bonnet D., Thome M., Tschopp J.,  
 RA FEBS Lett. 505:199-199(2001).  
 CC -1- FUNCTION: Activates NF-kappaB via Bcl10 and IKK. Stimulates the phosphorylation of Bcl10.  
 CC -1- SUBUNIT: CARD11 and Bcl10 bind to each other by CARD-CARD interaction.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- TISSUE SPECIFICITY: Detected in adult peripheral blood leukocytes, thymus, spleen and liver. Also found in promyelocytic leukemia HL-60 cells, chronic myelogenous leukemia K562 cells, Burkitt's lymphoma Raji cells and colorectal adenocarcinoma SW60 cells. Not detected in HeLa S3, Molt-4, A549 and G431 cells.  
 CC -1- SIMILARITY: Contains 1 CARD domain.  
 CC -1- SIMILARITY: Contains 1 PDZ/DHR domain.  
 CC -1- SIMILARITY: Contains 1 guanylate kinase-like domain.  
 CC -1- CAUTION: Supposed to contain a SH3 domain which is not detected by

PROSITE, Pfam or SMART.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; AF322641; AAC53402.1; -.  
 CC Genew; HGNC:16393; CARD11.  
 CC MIM: 607210; -.  
 DR GO; GO:0005624; C:membrane fraction; NAS.  
 DR GO; GO:0004384; F:membrane-associated guanylate kinase; NAS.  
 DR GO; GO:0005515; F:protein binding activity; IPI.  
 DR GO; GO:0007250; P:activation of NF-kappaB-inducing kinase; NAS.  
 DR InterPro; IPR001315; CARD.  
 DR InterPro; IPR000619; Guanylate\_kin.  
 DR InterPro; IPR001478; PDZ.  
 DR SMART; SMO0228; PDZ; 1.  
 DR PROSITE; PS50209; CARD; 1.  
 DR PROSITE; PS00856; GUANYLATE\_KINASE\_1; FALSE\_NEG.  
 DR PROSITE; PS50052; GUANYLATE\_KINASE\_2; FALSE\_NEG.  
 DR PROSITE; PS50106; PDZ; FALSE\_NEG.  
 KW Coiled coil.  
 FT DOMAIN 11 103 CARD.  
 FT DOMAIN 123 442 COILED COIL (POTENTIAL).  
 FT DOMAIN 673 748 PDZ.  
 FT DOMAIN 966 1133 GUANYLATE KINASE.  
 FT CONFLICT 808 808 P -> L (IN REF. 2).  
 SQ SEQUENCE 1147 AA; 132641 MW; 913A4B015D2B36CC CRC64;

Query Match 37.6%; Score 270.5; DB 1; Length 1147;  
 Best Local Similarity 46.6%; Pred. No. 5.1e-19;  
 Matches 54; Conservative 23; Mismatches 33; Indels 1; Gaps 1;

QY 15 DEBTLMEMESHRIRVICRCPRLTPYLRQAKVLCQDEEVLHSPRLTNSAMRAGHLL 74  
 DB 11 EBDALMEVENCNRMLSLSYINPAKLTIPYLRQCKYIDGDEDEEVANAPVLPKINRAGRL 70

QY 75 DLKTRGNKGAIAFESLKFHNPVYTLVTGLQPDVDFSNFSGSSDPDGL 125  
 DB 71 DILHTKGQGVVFLSHLFYPRILYKLVTKSEPTTRFSTIVEEG-HEGL 120

RESULT 4  
 CARA\_MOUSE STANDARD; PRT; 1021 AA.  
 ID CARA\_MOUSE STANDARD; PRT; 1021 AA.  
 AC P58660;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Caspase recruitment domain protein 10 (Bcl10-interacting MAGUK protein 1) (Bim1).  
 GN CARD10 OR BIM1.  
 OS Mus musculus (mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21391892; PubMed=11387339;  
 RA McAllister-Lucas L.M., Inohara N., Lucas P.C., Ruland J., Benito A.,  
 RA Li Q., Chen S., Chen F.F., Yamaoka S., Verma I.M., Mak T.W.,  
 RA Nunez G.,  
 RT "Bim1, a MAGUK family member linking protein kinase C activation to Bcl10-mediated NF-kappa B induction.";  
 RT J. Biol. Chem. 276:30589-30597(2001).  
 CC -1- FUNCTION: Activates NF-kappaB via Bcl10 and IKK.  
 CC -1- SUBUNIT: CARD10 and Bcl10 bind to each other by CARD-CARD interaction. They both participate in a complex with MALT1, where MALT1 binds to Bcl10.

CC -1- TISSUE SPECIFICITY: Highly expressed in kidney, heart followed by  
 CC brain, lung, liver, skeletal muscle and testis.  
 CC -1- SIMILARITY: Contains 1 CARD domain.  
 CC -1- CAUTION: Supposed to contain a SH3, a PDZ and a guanylate kinase-  
 CC like domain. But none of these 3 domains are detected by PROSITE,  
 CC Pfam or SMART.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF363456; AK60136.1; --  
 CC MGD: MGI:2146012; Card10.  
 CC DR GO: GO:0005737; Cytoplasm; ISS.  
 CC DR GO: GO:0005515; Protein binding activity; ISS.  
 CC DR GO: GO:0003015; Receptor signaling complex scaffold protein; ISS.  
 CC DR GO: GO:0007250; Phosphorylation of NF-kappa-inducing kinase; ISS.  
 CC DR GO: GO:0004651; Protein complex assembly; ISS.  
 CC DR InterPro: IPR001315; CARD.  
 CC DR PROSITE: PS50209; CARD; 1.  
 CC KM Coiled coil.  
 CC FT DOMAIN 23 115 CARD  
 CC FT DOMAIN 138 450 COILED COIL (POTENTIAL).  
 CC FT DOMAIN 558 565 POLY-SER.  
 CC SQ SEQUENCE 1021 AA; 114413 MW; 4811A09BDBF792C CRC64;  
 CC -----  
 CC Query Match 35.0%; Score 252; DB 1; Length 1021;  
 CC Best Local Similarity 51.1%; Pred. No. 3, 1e-17;  
 CC Matches 48; Conservative 20; Mismatches 26; Indels 0; Gaps 0;  
 CC -----  
 CC QY 15 DEETLMMESRRHPIVRCICPSRLPYLRQAKVLCQDESEVLSPLTNSAKRAGILL 74  
 CC Db 23 EEDALMEREGVRHRLTRALNPAKLTPIIRGCRVLEDEDESEVLSYFPCRAKRTGLI 82  
 CC QY 75 DLKTRGNKGAIAFLPSIKFNHPDVYITVLTQIP 108  
 CC Db 83 DILRCRKGKGFAPFLDALFEYFPHFTLLTQGP 116  
 CC -----  
 CC RESULT 5  
 CC CARD HUMAN STANDARD; PRT; 1032 AA.  
 CC ID CARA HUMAN  
 CC AC Q9BWT7; O9UGR5; O9Y3H0; PRT; 1032 AA.  
 CC DT 28-FEB-2003 (Rel. 41, Created)  
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 CC DE Caspase recruitment domain protein 10 (CARD-containing MAGUK protein  
 CC 3) (Carda 3).  
 CC GN CARD10 OR CARMA3.  
 CC OS Homo sapiens (Human).  
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 CC OX NCBI\_TaxID=9606;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RX MEDLINE=21292987; PubMed=11259443;  
 CC RA Wang L., Guo Y., Huang W.-U., Ke X., Poyet J.-L., Manji G.A.,  
 CC RA Merlam S., Glucksmann M.A., Distefano P.S., Alnemri B.S., Bertin J.;  
 CC RT "CARD10 is a novel caspase recruitment domain/membrane-associated  
 CC RT guanylate kinase family member that interacts with Bcl10 and activates  
 CC RT NF-kappa B.";  
 CC RT J. Biol. Chem. 276:21405-21409 (2001).  
 CC RN [2]  
 CC RP SEQUENCE FROM N.A.  
 CC RX MEDLINE=21255653; PubMed=11356195;  
 CC RA Gaide O., Marinton F., Michieu O., Bonnet D., Thome M., Tschopp J.;  
 CC RT "Catalin, a CARD-containing binding partner of Bcl10, induces Bcl10  
 CC RT phosphorylation and NF-kappaB activation.";  
 CC RT FEBS Lett. 496:121-127 (2001).  
 CC -----

CC [3]  
 CC RP ERRATUM.  
 CC RA Gaide O., Marinton F., Michieu O., Bonnet D., Thome M., Tschopp J.;  
 CC RT FEBS Lett. 505:198-198 (2001).  
 CC RN [4]  
 CC RP SEQUENCE FROM N.A.  
 CC RX MEDLINE=20057165; PubMed=10591208;  
 CC RA Dunham I., Hunt A.R., Collins J.E., Bruskiewicz R., Beare D.M.,  
 CC RA Bagdely C., Bailey J., Bartow K.F., Bates K.N., Beasly A.K.,  
 CC RA Bird C.P., Blakey S.E., Bridgman A.M., Buck D., Burgess J.,  
 CC RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,  
 CC RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,  
 CC RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,  
 CC RA Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,  
 CC RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,  
 CC RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,  
 CC RA Hall R.E., Hall-Tamlyn G., Heathcote R.M., Ho S., Holmes S.,  
 CC RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,  
 CC RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,  
 CC RA McLaren I.D., Mahegani-Mohammadi M., Matthews L.H., McCann O.T.,  
 CC RA McElay J., McLaren S., McMurtry A.A., Milne S.A., Mortimore B.J.,  
 CC RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,  
 CC RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,  
 CC RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,  
 CC RA Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,  
 CC RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,  
 CC RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,  
 CC RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,  
 CC RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,  
 CC RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Matsuyama S.,  
 CC RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,  
 CC RA Dorman A., Fang F., Fan Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,  
 CC RA Lewis J., Lewis S., Lin S.-P., Loh P., Majaj E., Nguyen T., Pan H.,  
 CC RA Pham S., Qi S., Qian Y., Ray L., Ren O., Shaili S., Song L.,  
 CC RA Zhang Q., Wang Y., Wang Z., White J., Williamson D., Wu H., Yao Z.,  
 CC RA Zhan M., Zhang G., Chisoe S., Murray J., Miller N., Mink P.,  
 CC RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,  
 CC RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,  
 CC RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Roloff T.,  
 CC RA Scheet P., Walker C., Mamsley A., Woldmann P., Pepin K., Nelson R.,  
 CC RA Korfi I., Bedell J.A., Hillier L., Wards E., Waterson R., Wilson R.,  
 CC RA Emanuel B.S., Shaikh T., Kurahashi H., Salita S., Budarf M.L.,  
 CC RA McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E., Belman L.,  
 CC RA Kim U.J., Shizuya H., Simon M.I., Dumanek J.P., Peyzard M., Kedra D.,  
 CC RA Seroussi E., Fransson I., Tapia I., Brudner C.E., O'Brien K.P.,  
 CC RA Wilkinson P., Boderech A., Hartman K., Hu X., Khan A.S., Lane L.,  
 CC RA Tliahun Y., Wright H.;  
 CC RT "The DNA sequence of human chromosome 22.";  
 CC RL Nature 402:489-495 (1999).  
 CC CC -1- SUBUNIT: CARD10 and Bcl10 bind to each other by CARD-CARD  
 CC interaction. They both participate in a complex with MALTI, where  
 CC MALTI binds to Bcl10 (by similarity).  
 CC CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC CC -1- TISSUE SPECIFICITY: Detected in adult heart, kidney and liver;  
 CC CC lower levels in intestine, placenta, muscle and lung. Also found  
 CC CC in fetal lung, liver and kidney.  
 CC CC -1- SIMILARITY: Contains 1 CARD domain.  
 CC CC -1- CAUTION: Supposed to contain a SH3, a PDZ and a guanylate kinase-  
 CC CC like domain. But none of these 3 domains are detected by PROSITE,  
 CC CC Pfam or SMART.  
 CC CC -1- CAUTION: Ref. 4 sequence differs from that shown due to various  
 CC CC gene identification problems.  
 CC CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

```

DR EMBL; AY028896; AAK26165.1; -.
DR EMBL; AY029228; AAK54454.1; -.
DR EMBL; AL049851; CAB63075.1; ALT_SEQ.
DR EMBL; AL049851; CAB63076.1; ALT_SEQ.
DR EMBL; AL022315; CAB42832.1; ALT_SEQ.
DR Genew; HGNC:16422; CARD10.
DR MIM; 607209; -.
DR GO; GO:0005737; C:cytoplasm; NAS.
DR GO; GO:0005515; F:protein binding activity; IPT.
DR GO; GO:0030159; F:receptor signaling complex scaffold protein. .; NAS.
DR GO; GO:0007250; P:activation of NF-kappa-inducing kinase; IDA.
DR GO; GO:0006461; P:protein complex assembly; NAS.
DR InterPro; IPR001315; CARD.
DR PROSITE; PS50209; CARD; 1.
KM Coiled coil.
FT DOMAIN 23 115 CARD.
FT DOMAIN 138 456 COILED COIL (POTENTIAL).
FT DOMAIN 567 574 POLY-SER.
FT CONFLICT 289 289 O -> R (IN REF. 4).
FT CONFLICT 917 917 K -> Q (IN REF. 4; CAB63075).
FT CONFLICT 932 932 R -> L (IN REF. 4).
SQ SEQUENCE 1032 AA; 115946 MW; 8377319AB82A0949 CRC64;

Query Match 34.2%; Score 246; DB 1; Length 1032;
Best Local Similarity 46.9%; Pred. No. 1.3e-16;
Matches 46; Conservative 21; Mismatches 27; Indels 0; Gaps 0;

QY 15 DEFTLMMESHRRHIVRCICPSRLTPYLRQAKVLCQDDEEVYLSPRITNSMRAQHLL 74
DB 23 EDLALMERIBGVKRLRALNPAKLTLYLRQCRVIDQDEEVYLSYTRFCRVNRTGRLL 82
QY 75 DLKTRGNKGAIAFLSLKFNPDVYTLVTGLOP 108
DB 83 DILRCRGVGEATVLELFYFPHFTLITNGEP 116

RESULT 6
CAR9_HUMAN STANDARD; PRT; 536 AA.
AC OSEPTO.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Caspase recruitment domain protein 9 (CARD9).
GN CARD9.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=20576268; PubMed=11053425;
RA Bertin J., Guo Y., Wang L., Srinivasula S.M., Jacobson M.D.,
RA Poyet J.-L., Meriam S., Du M.-Q., Dyer M.J.S., Robison K.E.,
RA Distefano P.S., Alnemri E.S.;
RA "CARD9 is a novel caspase recruitment domain-containing protein that
RA interacts with Bcl10/CLAP and activates NF-kappa B."
RT J. Biol. Chem. 275:41082-41086(2000).
RL -1- FUNCTION: Activates NF-kappaB via Bcl10 (By similarity).
RL -1- SUBUNIT: Self-associates. CARD9 and Bcl10 bind to each other by
RL -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
RL -1- SIMILARITY: Contains 1 CARD domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
```

```

DR EMBL; AF311288; AAG28791.1; -.
DR InterPro; IPR001315; CARD.
DR PROSITE; PS50209; CARD; 1.
KM Coiled coil.
FT DOMAIN 6 98 CARD.
FT DOMAIN 117 277 COILED COIL (POTENTIAL).
FT DOMAIN 303 420 COILED COIL (POTENTIAL).
SQ SEQUENCE 536 AA; 62631 MW; 6F33089CB7E6BAC9 CRC64;

Query Match 31.0%; Score 223.5; DB 1; Length 536;
Best Local Similarity 45.4%; Pred. No. 1e-14;
Matches 49; Conservative 18; Mismatches 38; Indels 3; Gaps 1;

QY 15 DEFTLMMESHRRHIVRCICPSRLTPYLRQAKVLCQDDEEVYLSPRITNSMRAQHLL 74
DB 6 NDECSWALSFPVYKLSIVIDPSRTIPYLRQCKVNPDEBQVLDPNIVIRKRGVLL 65
QY 75 DLKTRGNKGAIAFLSLKFNPDVYTLVTGLOPVDPS--NFGES 119
DB 66 DILQTRGHGVYAFLESLKLYPQLYRKVTGEPARVPSMTIDASGES 113

RESULT 7
CAR9_HUMAN STANDARD; PRT; 536 AA.
AC Q9H257; Q9H854;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Caspase recruitment domain protein 9 (CARD9).
GN CARD9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20576268; PubMed=11053425;
RA Bertin J., Guo Y., Wang L., Srinivasula S.M., Jacobson M.D.,
RA Poyet J.-L., Meriam S., Du M.-Q., Dyer M.J.S., Robison K.E.,
RA Distefano P.S., Alnemri E.S.;
RA "CARD9 is a novel caspase recruitment domain-containing protein that
RA interacts with Bcl10/CLAP and activates NF-kappa B."
RT J. Biol. Chem. 275:41082-41086(2000).
RL [2]
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Retinoblastoma;
RA Isegaki T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagatsuna M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto K., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y.,
RA Ninomiya K., Iwamoto T.;
RA "NEBD human cDNA sequencing project."
RT Submitted (AUG-2000) to the EMBL/GenBank/DBSJ databases.
RN [3]
RP SEQUENCE OF 1-492 FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RC TISSUE=Muscle;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheefter C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Slaughter M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein W.J., Uediri T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mollath S.J.,
RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fehy J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
```

RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W.,  
 RA Butlerfield Y.S.N., Krzyzanski M.I., Skalska U., Smailus D.E.,  
 RA Scheraga A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- FUNCTION: Activates NF-kappaB via Bcl10.  
 CC -1- SUBUNIT: Self-associates. CARD9 and Bcl10 bind to each other by  
 CC CARD-CARD interaction.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- TISSUE SPECIFICITY: Highly expressed in spleen. Also detected in  
 CC liver, placenta, lung, peripheral blood leukocytes and in brain.  
 CC -1- SIMILARITY: Contains 1 CARD domain.  
 CC -1- CAUTION: Ref.2 sequence differs from that shown due to a  
 CC frameshift in position 360.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL; AF31287; AAG28790.1; -;  
 DR EMBL; AK024001; BAB14766.1; ALT\_FRAME.  
 DR EMBL; BC008877; AAH08877.1; -;  
 DR GeneW; HGNC:16391; CARD9.  
 DR MIM; 607212; -;  
 DR InterPro; IPR001315; CARD.  
 DR PROSITE; PS50209; CARD; 1.  
 KM Coiled coil.  
 FT DOMAIN 6 98 CARD.  
 FT 117 277 COILED COIL (POTENTIAL).  
 FT DOMAIN 332 419 COILED COIL (POTENTIAL).  
 FT 12 12 N -> S (IN REF. 3).  
 FT CONFLICT 482 492 LSGEPPER -> PAGRGIGAVC (IN REF. 3).  
 FT CONFLICT 536 AA; 62267 MW; 6EB1835315B83DE5 CR664;  
 SQ SEQUENCE  
 Query Match 29.9%; Score 215.5; DB 1; Length 536;  
 Best Local Similarity 44.4%; Pred. No. 6.4e-14;  
 Matches 48; Conservative 17; Mismatches 40; Indels 3; Gaps 1;  
 QY 15 DEETLWEMSHRHRIVRCIPSRRLTPYLRQAKVLCQDDEEVLHSPRTNSAPRAGHLL 74  
 DB 6 NDECNVAVGEFVTLTSVDESRITPYLRQCKVLPDDEEQLSPNIVIRKRVGL 65  
 QY 75 DLKTKGKAGATAFESLKFHNDVYTLVTGLQDPVDS--NFGGS 119  
 DB 66 DILQRTGHKGYAFLESLELYYPOLKYKVGKPARVFSMIIDAGS 113  
 RESULT 8  
 BCL1\_HUMAN STANDARD; PRT; 233 AA.  
 ID BCL1\_HUMAN STANDARD; PRT; 233 AA.  
 AC 095999;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE B cell lymphoma/leukemia 10 (B-cell CLL/lymphoma 10) (Bcl-10) (CED-  
 DE 3/ICH-1 prodomain homologous E10-like regulator) (CIPEP) (CARD-  
 DE containing molecule enhancing NF-kappaB) (cellular homolog of VCARMEN)  
 DE (VCARMEN) (Mammalian CARD-containing adapter molecule E10) (ME10)  
 DE (cellular-E10) (C-E10) (CARD-like apoptotic protein) (hCLAP).  
 GN BCL10 OR CIPEP OR CLAP.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RN SEQUENCE FROM N.A., VARIANT FOLLICULAR LYMPHOMA GLU-210 DEL, VARIANT  
 RP MESOTHELIOMA ILE-52, AND VARIANTS GERM CELL TUMOR GLY-58 AND PRB-219.

RC TISSUE=Lymphoma;  
 RX MEDLINE=99142601; PubMed=9989495;  
 RA Willis T.G., Jadayel D.M., Du M.-Q., Peng H., Perry A.R.,  
 RA Abdul-Rauf M., Price H., Karan L., Matekodom O., Mlodarska I.,  
 RA Par L., Crook T., Hamoudi R., Isaacson P., Dyer M.J.S.;  
 RT "Bcl10 is involved in t(1;14)(p22;q32) of MALT B cell lymphoma and  
 RT mutated in multiple tumor types.";  
 RL Cell 96:35-45(1999).  
 RN [2]  
 RN SEQUENCE FROM N.A., AND MUTAGENESIS OF LEU-41 AND GLY-78.  
 RX MEDLINE=99214545; PubMed=10187770;  
 RA Koseki T., Inohara N., Chen S., Carrio R., Merino J., Hottiger M.O.,  
 RA Nabel G.J., Nunez G.;  
 RT "CIPEP, a novel NF-kappaB-activating protein containing a caspase  
 RT recruitment domain with homology to Herpesvirus-2 protein E10.";  
 RL J. Biol. Chem. 274:9955-9961(1999).  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=99214546; PubMed=10187771;  
 RA Thome M., Martinon F., Hofmann K., Rubio V., Steiner V., Schneider P.,  
 RA Mattmann C., Tschopp J.;  
 RT "Equine herpesvirus-2 E10 gene product, but not its cellular  
 RT homologue, activates NF-kappaB transcription factor and c-Jun  
 RT N-terminal kinase.";  
 RL J. Biol. Chem. 274:9962-9968(1999).  
 RN [4]  
 RN SEQUENCE FROM N.A., AND MUTAGENESIS OF LEU-28; LEU-41; ILE-46; LEU-47;  
 RP GLU-53 AND ILE-55.  
 RX MEDLINE=99214590; PubMed=10187815;  
 RA Yan M., Lee J., Schilbach S., Goddard A., Dixit V.M.;  
 RT "mE10, a novel caspase recruitment domain-containing proapoptotic  
 RT molecule";  
 RL J. Biol. Chem. 274:10287-10292(1999).  
 RN [5]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=99292766; PubMed=10364242;  
 RA Srinivasula S.M., Ahmed M., Lin J.-H., Poyet J.-L.,  
 RA Fernandes-Alnemri T., Tsichlis P.N., Alnemri E.S.;  
 RT "CLAP, a novel caspase recruitment domain-containing protein in the  
 RT tumor necrosis factor receptor pathway, regulates NF-kappaB  
 RT activation and apoptosis";  
 RL J. Biol. Chem. 274:17946-17954(1999).  
 RN [6]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=Spleen;  
 RX MEDLINE=99329013; PubMed=10400625;  
 RA Costanzo A., Guet C., Vitco P.;  
 RT "C-E10 is a caspase-recruiting domain-containing protein that  
 RT interacts with components of death receptors signaling pathway and  
 RT activates nuclear factor-kappaB";  
 RL J. Biol. Chem. 274:20127-20132(1999).  
 RN [7]  
 RN SEQUENCE FROM N.A., AND VARIANTS MALT LYMPHOMA SER-5; GLU-16; GLU-31;  
 RP ARG-57; LYS-64; GLU-101; PRO-134; ALA-168; SER-174; GLU-213 AND  
 RP ILE-230.  
 RX MEDLINE=99251581; PubMed=10319863;  
 RA Zhang Q., Siebert R., Yan M., Hinzmann B., Gui X., Xue L.,  
 RA Rakeschew K.M., Naeye C.W., Beckmann G., Weisenburger D.D.,  
 RA Sanger W.G., Nowotny H., Vesely M., Callet-Bauchu E., Salles G.,  
 RA Dixit V.M., Rosenthal A., Schlegelberger B., Morris S.W.;  
 RT "Inactivating mutations and overexpression of BCL10, a caspase  
 RT recruitment domain-containing gene, in MALT lymphoma with  
 RT t(1;14)(p22;q32).";  
 RL Nat. Genet. 22:63-68(1999).  
 RN [8]  
 RN PHOSPHORYLATION.  
 RX MEDLINE=21359551; PubMed=11466612;  
 RA Yui D., Yoneda T., Oono K., Katayama T., Imai K., Tohyama M.;  
 RT "Interchangeable binding of Bcl10 to TRAF2 and CIAPs regulates  
 RT apoptosis signaling";  
 RL Oncogene 20:4317-4323(2001).  
 RN [9]  
 RN VARIANTS MESOTHELIOMA SER-5; GLN-45; GLN-58; SER-93; VAL-153; GLU-213

RP	AND PHE-218.
RX	MEDLINE=99308628; PubMed=10380921;
RA	Apostolou S., de Rianzo A., Worthy S.S., Jhanwar S.C., Testa J.R.;
RT	"Absence of BCL10 mutations in human malignant mesothelioma.";
RL	Cell 97:684-686(1999).
CC	-1- FUNCTION: Promotes apoptosis, pro-caspase-9 maturation and activation of NF-kappaB via NIK and IKK. May be an adapter protein between upstream TNFR1-TRADD-RIP complex and the downstream NIK-IKK-IRAP complex.
CC	-1- SUBUNIT: Self associates by CARD-CARD interaction and forms a tight complex with MAL1. Interacts with other CARD-proteins such as CARD9, CARD10, CARD11 and CARD14. Binds caspase-9 with its C-terminal domain. Interacts with TRAF2 and BirC2/C-IP2.
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic. Appears to have a perinuclear, compact and filamentous pattern of expression. Also found in the nucleus of several types of tumor cells.
CC	-1- TISSUE SPECIFICITY: Ubiquitous.
CC	-1- PTM: Phosphorylated. Phosphorylation results in dissociation from TRAF2 and binding to BirC2/C-IP2.
CC	-1- DISEASE: Involved in a t(1;14)(p22;q32) chromosomal translocation recurrent in low-grade MALT lymphoma (mucosa-associated lymphoid tissue). Although the Bcl10/IgH translocation leaves the coding region of Bcl10 intact, frequent Bcl10 mutations could be attributed to the Ig somatic hypermutation mechanism resulting in nucleotide transitions.
CC	-1- DISEASE: Defects in BCL10 are involved in various types of cancer.
CC	-1- SIMILARITY: Contains 1 CARD domain.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; AJ006288; CA006955.1; -
DR	EMBL; AF057700; AAD15800.1; -
DR	EMBL; AF100338; AAD16428.1; -
DR	EMBL; AF127386; AAD32597.1; -
DR	EMBL; AF134395; AAD39147.1; -
DR	EMBL; AF105066; AAP06894.1; -
DR	EMBL; AF082283; AAC99767.1; -
DR	EMBL; AF097732; AAD24918.1; -
DR	GeneW; HGNC; 989; BCL10.
DR	MTM; 603517; -
DR	GO; GO:0005515; F:protein binding activity; IPI.
DR	GO; GO:0008181; F:tumor suppressor; TAS.
DR	GO; GO:0006917; P:induction of apoptosis; TAS.
DR	InterPro; IPR01315; CARD.
DR	Pfam; PF00619; CARD_1.
DR	SMART; SM00114; CARD; 1.
DR	PROSITE; PS50209; CARD; 1.
KM	Apoptosis; Chromosomal translocation; Anti-oncogene; Phosphorylation;
KM	Disease mutation.
FT	DOMAIN
FT	VARIANT
FT	13 101
FT	5 5
FT	CARD.
FT	A->S (IN MALT LYMPHOMA AND MESOTHELIOA).
FT	/FtId=VAR_013208.
FT	V->E (IN MALT LYMPHOMA).
FT	/FtId=VAR_013309.
FT	K->E (IN MALT LYMPHOMA).
FT	/FtId=VAR_013210.
FT	K->Q (IN MESOTHELIOA).
FT	/FtId=VAR_013211.
FT	T->I (IN MESOTHELIOA).
FT	/FtId=VAR_013212.
FT	C->R (IN MALT LYMPHOMA).
FT	/FtId=VAR_013213.
FT	R->G (IN GERM CELL TUMOR).
FT	/FtId=VAR_013214.
FT	R->Q (IN MESOTHELIOA).
FT	/FtId=VAR_013215.
FT	VARIANT
FT	58 58

[illegible]



RA Sato N., Gomi F., Morihara T., Mori Y., Miyoshi K., Hitomi U.,  
 RA Ugawa S., Yamada S., Okabe M., Tohyama M.,  
 RT "Regulatory mechanisms of TRAF2-mediated signal transduction by Bcl10,  
 RT a MyD88 lymphoma-associated protein.";  
 RL J. Biol. Chem. 275:11114-11120(2000).  
 CC -1- FUNCTION: Promotes apoptosis, pro-caspase-9 maturation and  
 CC activation of NF-kappaB via NIK and IKK. May be an adapter protein  
 CC between upstream TNFR1-TRADD-RIP complex and the downstream NIK-  
 CC IKK-IKAP complex (By similarity).  
 CC -1- SUBUNIT: Self-associates by CARD-CARD interaction and interacts  
 CC with other CARD-proteins such as CARD9, CARD10, CARD11 and CARD14.  
 CC Binds caspase-9 with its C-terminal domain (By similarity).  
 CC Interacts with TRAF2 and BIRC2/c-IAP2.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- PTM: Phosphorylated (By similarity).  
 CC -1- SIMILARITY: Contains 1 CARD domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AB016069; BAA88822.1; -  
 DR InterPro: IPR001315; CARD.  
 DR Pfam: PF00619; CARD; 1.  
 DR SMART: SM00114; CARD; 1.  
 DR PROSITE: PS50209; CARD; 1.  
 KM Apoptosis; Anti-oncogene; Phosphorylation.  
 FT DOMAIN 13 CARD  
 SQ SEQUENCE 233 AA; 25999 MW; B43274B4B825FC7D CRC64;  
 Query Match 11.2%; Score 81; DB 1; Length 233;  
 Best Local Similarity 36.5%; Pred. No. 0.63;  
 Matches 31; Conservative 14; Mismatches 32; Indels 8; Gaps 4;  
 QY 13 ALDEBTWEN---MESHRIYRCIPSLTYLRYLQAKYLCLDDEEVLHSERLTNSM 68  
 DB 7 SLNEEDLVEYKDALERLRYLCEKILAEHRHFDHLRAKILSRDETEI--SCR-TSSRK 63  
 QY 69 RAGHLDLILTRGKNGAIFLSLX 93  
 DB 64 RAGKLDLYIQENPR-GIDTLVESIR 87  
 RESULT 10  
 UVRC METTH STANDARD; PRT; 579 AA.  
 ID UVRC METTH STANDARD; PRT; 579 AA.  
 AC 026541;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE UVRAC System protein C (UVRC protein) (Exonuclease A3C subunit C).  
 GN UVRAC OR MTH441.  
 OS Methanobacterium thermoautotrophicum.  
 CC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;  
 CC Methanobacteriaceae; Methanobacter.  
 CC NCBI\_TaxID=187420;  
 CX [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=Delta H;  
 RX MEDLINE=98037514; PubMed=9371463;  
 RA Smith D.R., Doucette-Stamm L.A., Delonghery C., Lee H.-M., Dubois J.,  
 RA Aldrege T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,  
 RA Harrison D., Hoang L., Keagle P., Lamm W., Potlher B., Qiu D.,  
 RA Spadator R., Vycare R., Wang Y., Wierzbowski J., Gibson R.,  
 RA Jivan N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,  
 RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,  
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.,  
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
 RT delta: functional analysis and comparative genomics.";

RL J. Bacteriol. 179:7135-7155(1997).  
 CC -1- FUNCTION: The UVRAC repair system catalyzes the recognition and  
 CC processing of DNA lesions. UVRC both incises the 5' and 3' sides  
 CC of the lesion. The N-terminal half is responsible for the 3'  
 CC incision and the C-terminal half is responsible for the 5'  
 CC incision (By similarity).  
 CC -1- SUBUNIT: Interacts with UVRB in an incision complex (By  
 CC similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: Belongs to the UVRC family.  
 CC -1- SIMILARITY: Contains 1 UVR domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AE000828; AAB84947.1; -  
 DR PIR: F69157;  
 DR HSRP: P07023; IE52.  
 DR HAMAP: MF\_00203; -; 1.  
 DR InterPro: IPR000445; HHN.  
 DR InterPro: IPR003583; HHN 1.  
 DR InterPro: IPR001943; UVRB/C.  
 DR InterPro: IPR004791; UVRB.  
 DR InterPro: IPR001162; UVR\_C.  
 DR InterPro: IPR003050; UVR\_N.  
 DR Pfam: PF01541; Excl\_endo\_N; 1.  
 DR Pfam: PF00633; HHN; 2.  
 DR Pfam: PF02151; UVR; 1.  
 DR PRODOM: PD008670; UVR\_C; 1.  
 DR SMART: SM00465; GYFC; 1.  
 DR TIGRFAMs: TIGR0194; UVR\_C; 1.  
 DR PROSITE: PS50164; UVR\_C; 1.  
 DR PROSITE: PS50165; UVR\_C; 2; 1.  
 KM SOS response; Excision nuclease; DNA repair; DNA recombination;  
 KM DNA excision; Complete proteome.  
 FT DOMAIN 193 228 UVR.  
 SQ SEQUENCE 579 AA; 66293 MW; 83D3D7B8F9E3A68 CRC64;  
 Query Match 11.2%; Score 81; DB 1; Length 579;  
 Best Local Similarity 27.6%; Pred. No. 1.9;  
 Matches 43; Conservative 26; Mismatches 53; Indels 34; Gaps 9;  
 QY 7 RDSALTADDELT-----WEMES---HRRIRYRCIPSLTYLRQ---AVYLQOLD 53  
 DB 262 RDGKITGDPDFILRGSAPRTIELFAFKQYAIRRVSEILTYPVDDGYIAMLSLR 321  
 QY 54 EEEV-LHSP-----RLTNSAMRAGHLLDLKTRGKNGAIFLSLKHFNPDVYTLVGL 106  
 DB 322 GEVYKHSPEGAGRLINIMKASVILKQKGRVRLDLKQDKL--PEIRRMGL 379  
 QY 107 QPDVPSNFSGSSD-----FDG--LAGTSRNL 134  
 DB 360 -----DISNIAGESATGSVAVPIDKPSGSGYRRYRI 411  
 RESULT 11  
 AMPN RABIT STANDARD; PRT; 965 AA.  
 ID AMPN RABIT STANDARD; PRT; 965 AA.  
 AC P15541;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Aminopeptidase N (EC 3.4.11.2) (rBAPN) (Alanyl aminopeptidase)  
 DE (Microsomal aminopeptidase) (Aminopeptidase M) (Leukemia antigen  
 DE CD13).  
 GN ANPBP.



OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 NC NCB1\_TaxID=9986;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RX MEDLINE=94099942; Pubmed=7903857;  
 RT "Complete sequence of rabbit kidney aminopeptidase N and mRNA localization in rabbit kidney by in situ hybridization.";  
 RL Biochem. Cell Biol. 71:278-287(1993).  
 RP [2]  
 RP SEQUENCE OF 5-18.  
 RX MEDLINE=82113673; Pubmed=6120002;  
 RA Peracchi H., Maroux S., Bonicel J., Desnuelle P.;  
 RT "The amino acid sequence of the hydrophobic anchor of rabbit intestinal brush border aminopeptidase N.";  
 RL Biochim. Biophys. Acta 684:133-136(1982).  
 RN [3]  
 RN SEQUENCE OF 177-965 FROM N.A.  
 RX MEDLINE=90092508; Pubmed=2574692;  
 RA Noren O., Dabelsteen E., Hoeyer P.E., Olsen J., Sjoestrom H.,  
 RT "Onset of transcription of the aminopeptidase N (leukemia antigen CD 13) gene at the crypt/villus transition zone during rabbit enterocyte differentiation.";  
 RL FEBS Lett. 259:107-112(1989).  
 CC -1- FUNCTION: Broad specificity aminopeptidase. Plays a role in the final digestion of peptides generated from hydrolysis of proteins by gastric and pancreatic proteases. May be involved in the metabolism of regulatory peptides of diverse cell types and in the cleavage of peptides bound to major histocompatibility complex class II molecules of antigen presenting cells. May have a role in angiogenesis (by similarity).  
 CC -1- CATALYTIC ACTIVITY: Release of an N-terminal amino acid, Xaa-|-Xbb-from a peptide, amide or arylamide. Xaa is preferably Ala, but may be most amino acids including Pro (slow action). When a terminal hydrophobic residue is followed by a prolyl residue, the two may be released as an intact Xaa-Pro dipeptide.  
 CC -1- COFACTOR: Binds 1 zinc ion (by similarity).  
 CC -1- SUBUNIT: Homodimer (by similarity).  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein.  
 CC -1- PFM: Sulfated (by similarity).  
 CC -1- SIMILARITY: Belongs to peptidase family M1.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: S68687; AAB29334.1; -;  
 DR EMBL: X51508; CA35873.1; -;  
 DR MEROPS: M01.001; -;  
 DR InterPro: IPR001930; Ala\_peptase.  
 DR InterPro: IPR006025; Zn\_MTPeptide.  
 DR Pfam: PF01433; Peptidase\_M1; 1.  
 DR PRINTS: PR00756; ALADIPASE.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 KW Angiogenesis; Hydrolyase; Aminopeptidase; Metalloprotease; Zinc;  
 KW Signal-anchor; Transmembrane; Glycoprotein; Sulfation.  
 FT INIT MET 0  
 FT DOMAIN 1 7  
 FT TRANSMEM 8 31  
 FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 FT POTENTIAL.  
 FT CYTOSOLIC SER/THR-RICH JUNCTION.  
 FT DOMAIN 32 64  
 FT DOWNIN 65 965  
 FT METAL 383 384  
 FT ACT SITE 384 384  
 FT METAL 387 387  
 FT ZINC (CATALYTIC) (BY SIMILARITY).

FT METAL 406 406  
 FT ACT SITE 422 422  
 FT MOD\_RES 172 172  
 FT MOD\_RES 414 414  
 FT MOD\_RES 419 419  
 FT MOD\_RES 911 911  
 FT CARBOHYD 39 39  
 FT CARBOHYD 124 124  
 FT CARBOHYD 258 258  
 FT CARBOHYD 314 314  
 FT CARBOHYD 551 551  
 FT CARBOHYD 569 569  
 FT CARBOHYD 623 623  
 FT CARBOHYD 733 733  
 FT CARBOHYD 816 816  
 FT CONFLICT 9 9  
 FT CONFLICT 15 15  
 FT CONFLICT 209 209  
 FT CONFLICT 225 225  
 FT CONFLICT 233 233  
 FT CONFLICT 592 593  
 SQ SEQUENCE 965 AA; 109186 MW; 012D1895D457A96 CRC64;  
 Query Match 11.1%; Score 80; DB 1; Length 965;  
 Best Local Similarity 28.0%; Pred. No. 4.2;  
 Matches 44; Conservative 20; Mismatches 55; Indels 38; Gaps 10;  
 QY 9 SALTLDETLWMESHRIRIVRCIPSR---LTPYL-----RQAKVLCQ 51  
 Db 56 SPATLDONLWPK-----NRYLPKTLTIDSYNVVLRPLTSNSGLYFTGSSYTRFCQ 110  
 QY 52 -LDSEVLSHSPRLTNSARAGHLLDLTRGKN-GAIAFLSLKPHNDVY---TLVTG 105  
 Db 111 EATVNTIHSKTL-NYTTQGRVVLRGVRSQPPALSTELVELTVVHLQGLVAG 169  
 QY 106 LQPDVDFSNFSGESSDFDGLAGTSR-----NRLILV 136  
 Db 170 SQYEND-TQFGELLA--DDLGFYRSEYMEGNVRKV 203  
 RESULT 12  
 KC64 HUMAN STANDARD; PRT; 519 AA.  
 ID KC64 HUMAN  
 AC O8TDN1; Q96H24;  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 42, Last annotation update)  
 DE Potassium voltage-gated channel subfamily G member 4 (Potassium channel Kv6.3).  
 GN KCNG4 OR KCNG3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NC NCB1\_TaxID=9606;  
 RN [1]  
 RC SEQUENCE FROM N.A. (ISOFORM 1).  
 RX TISSUE=Brain;  
 RX MEDLINE=22056098; Pubmed=12060745;  
 RA Otschytsch N., Raes A., Van Hoorick D., Snyder D.J.;  
 RT "Obligatory heterotrimerization of three previously uncharacterized Kv channel alpha-subunits identified in the human genome.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:7986-7991(2002).  
 RP [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE=Muscle;  
 RX MEDLINE=22388257; Pubmed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Murusana K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein W.J., Uediri T.B., Toshlyuk S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Murry D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 CC -1- FUNCTION: Potassium channel subunit. Modulates channel activity by  
 CC shifting the threshold and the half-maximal activation to more  
 CC negative values.  
 CC -1- SUBUNIT: Heteromultimer with KCNB1, KCNC1 and KCNF1. Does not form  
 CC homomultimers.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Has to be  
 CC associated with KCNB1 or possibly another partner to get inserted  
 CC in the plasma membrane. Remains intracellular in the absence of  
 CC KCNB1.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=Q8TDN1-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q8TDN1-2; Sequence=VSP\_001029, VSP\_001030;  
 CC Note=No experimental confirmation available;  
 CC -1- TISSUE SPECIFICITY: Highly expressed in brain, and at lower levels  
 CC in liver, small intestine and colon.  
 CC -1- DOMAIN: The segment S4 is probably the voltage-sensor and is  
 CC characterized by a series of positively charged amino acids at  
 CC every third position.  
 CC -1- SIMILARITY: BELONGS TO THE VOLTAGE-GATED POTASSIUM CHANNEL FAMILY.  
 CC G SUBFAMILY.  
 CC  
 CC This SNR5S-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AF348984; AAH83911.1; -  
 CC EMBL: BC008969; AAH08969.1; -  
 CC Genew: HGNC:19697; KCNG4.  
 CC MIM: 607603; -  
 DR InterPro: IPR000210; BTB\_POZ.  
 DR InterPro: IPR005821; Ion\_trans.  
 DR InterPro: IPR001622; K+channel\_pore.  
 DR InterPro: IPR003131; K\_tetra.  
 DR InterPro: IPR005820; M+channel\_nlg.  
 DR Pfam: PF00520; Ion\_trans; 1.  
 DR Pfam: PF02214; K\_tetra; 1.  
 DR SMART: SMO0225; BTB; 1.  
 DR Transport: Ion transport; Ionic channel; Voltage-gated channel;  
 KM Potassium channel; Potassium transport; Potassium; Transmembrane;  
 KM Multigene family; Alternative splicing.  
 KW DOMAIN 1  
 FT TRANSMEM 119 139 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 259 279 SEGMENT S1 (POTENTIAL).  
 FT DOMAIN 280 289 SEGMENT S2 (POTENTIAL).  
 FT TRANSMEM 290 310 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 335 335 SEGMENT S3 (POTENTIAL).  
 FT TRANSMEM 335 355 SEGMENT S4 (POTENTIAL).  
 FT TRANSMEM 356 370 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 371 391 SEGMENT S5 (POTENTIAL).  
 FT TRANSMEM 406 426 SEGMENT H5 (PORE-FORMING) (POTENTIAL).  
 FT TRANSMEM 434 454 SEGMENT S6 (POTENTIAL).  
 FT DOMAIN 455 519 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 378 378 POLY-LEU.  
 FT VARSPIC 253 GECS -> VSGL (in isoform 2).

FT FT /FTId=VSP\_001029.  
 FT VARSPIC 257 519 Missing (in isoform 2).  
 FT FT /FTId=VSP\_001030.  
 SQ SEQUENCE 519 AA; 58979 MW; 55BBA355931AB0A4 CRC64;  
 Query Match 11.0%; Score 79; DB 1; Length 519;  
 Best Local Similarity 25.8%; Pred. No. 2.6;  
 Matches 33; Conservative 21; Mismatches 56; Indels 18; Gaps 5;  
 QY 3 ELCDSDALNALADBEITLWEMWESHRIHIVACIPSRITPYLRQAKYLCCQDEEYVHSPR 62  
 DB 137 EMC---ALSPQELAYWGIEBAH---LERC-CIRKILRLIELEBAKLRDVLKQR 188  
 QY 63 LT---NSMRAGHLDLILKTRGN-----GAIAPLESIKFNPVYTLVTLGQPDVDF 112  
 DB 189 ETRPAPHSRRWGLCMKRLKEMENPQSGLPGRVACSLTLPATYAVSLCVSTMPDLRA 248  
 QY 113 SNFSGESS 120  
 DB 249 EEDGEGCS 256  
 RESULT 13  
 ID BCLA\_MOUSE STANDARD; PRT; 233 AA.  
 AC 0920H7;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE B-cell lymphoma/leukemia 10 (B-cell CLL/lymphoma 10) (Bcl-10) (CED-3/ICH-1 prodomain homologous E10-like regulator) (mCIPER) (CARD-3(CARMEN) (Mammalian CARD-containing adapter molecule E10) (mE10) (Cellular-E10) (C-E10) (CARD-like apoptotic protein) (mCLAP).  
 DE BCL10 OR CIPER OR CLAP.  
 GN Mus musculus (Mouse).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99142601; PubMed=9989495;  
 RA Willis T.G., Jadayel D.M., Du M.-Q., Peng H., Perry A.R.,  
 RA Abdul-Rauf M., Price H., Kairan L., Mayekodunni O., Wlodarska I.,  
 RA Pan L., Crook T., Hamoudi R., Isaacson P., Dyer M.J.S.;  
 RT "Bcl10 is involved in t(1;14) (p22;q32) of MALT B cell lymphoma and  
 RT mutated in multiple tumor types.";  
 RL Cell 96:35-45 (1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99214545; PubMed=10187770;  
 RA Kosaki T., Inohara N., Chen S., Carrioe R., Merino J., Hottinger M.O.,  
 RA Nabel G.J., Nunez G.;  
 RT "CIPER, a novel NF-kappaB-activating protein containing a caspase  
 RT recruitment domain with homology to Herpesvirus-2 protein E10.";  
 RL J. Biol. Chem. 274:9955-9961 (1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99214546; PubMed=10187771;  
 RA Thome M., Martillon P., Hofmann K., Rubio V., Steiner V., Schneider P.,  
 RA Matman C., Tschopp J.;  
 RT "Equine herpesvirus-2 E10 gene product, but not its cellular  
 RT homologue, activates NF-kappaB transcription factor and c-Jun N-  
 RT terminal kinase.";  
 RL J. Biol. Chem. 274:9962-9968 (1999).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99214590; PubMed=10187815;  
 RA Yan M., Lee J., Schilbach S., Goddard A., Dixit V.M.;  
 RT "mE10, a novel caspase recruitment domain-containing proapoptotic  
 RT molecule.";  
 RL J. Biol. Chem. 274:10287-10292 (1999).

RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9292766; PubMed=10364242;  
 RA Srinivasula S.M., Ahmad M., Lin J.-H., Poyet J.-L.,  
 RA Fernandes-Alnemri T., Tsichlis P.N., Alnemri E.S.,  
 RT "cIAP, a novel caspase recruitment domain-containing protein in the  
 RT tumor necrosis factor receptor pathway, regulates NF-kappaB activation  
 RT and apoptosis.";  
 RL J. Biol. Chem. 274:17946-17954 (1999).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Straube R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Stenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marisina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,  
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Millaly S.J.,  
 RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gumaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butcherfield Y.S.N., Krzywinski M.I., Skala U., Smalins D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 CC -1- FUNCTION: Promotes apoptosis, pro-caspase-9 maturation and  
 CC activation of NF-kappaB via NIK and IKK. May be an adapter protein  
 CC between upstream TNFR1-TRADD-RIP complex and the downstream NIK-  
 CC IKK-IKAP complex (By similarity).  
 CC -1- SUBUNIT: Self-associates by CARD-CARD interaction and interacts  
 CC with other CARD-proteins such as CARD9, CARD10, CARD11 and CARD14.  
 CC -1- Binds caspase-9 with its C-terminal domain. Interacts with TRAF2  
 CC and BIRC2/C-1A2 (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- TISSUE SPECIFICITY: Highly expressed in heart, brain, spleen,  
 CC lung, liver, skeletal muscle, kidney and testis. Detected in  
 CC developing brain, olfactory epithelium, tongue, whisker follicles,  
 CC salivary gland, heart, lung, liver and intestinal epithelia of  
 CC stage 15 embryos.  
 CC -1- PTM: Phosphorylated (By similarity).  
 CC -1- SIMILARITY: Contains 1 CARD domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/annouce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 CC EMBL: AJ006289; CAA06956.1; -  
 CC EMBL: AF057701; AAD15801.1; -  
 CC EMBL: AF100339; AAD16429.1; -  
 CC EMBL: AF127387; AAD32598.1; -  
 CC EMBL: AF134396; AAD39148.1; -  
 CC EMBL: BC024379; AAH24379.1; -  
 CC MGD: MG11337994; Bcl10.  
 CC GO: GO:0006915; P:apoptosis; IDA.  
 CC GO: GO:0016066; P:humoral defense mechanism (sensu Vertebrata); IMP.  
 CC GO: GO:0016064; P:humoral defense mechanism (sensu Vertebrata); IMP.  
 CC InterPro: IPR001315; CARD.  
 CC Pfam: PF00619; CARD; 1.  
 CC SMART: SMO0114; CARD; 1.  
 CC PROSITE: PSS0209; CARD; 1.  
 CC Apoptosis; Anti-oncogene; Phosphorylation.

FT DOMAIN 13 101 CARD  
 SQ SEQUENCE 233 AA; 25948 MW; C0539BC97102DB8 CRC64;  
 Query Match 10.8%; Score 78; DB 1; Length 233;  
 Best Local Similarity 35.3%; Pred. No. 1.3;  
 Matches 30; Conservative 14; Mismatches 33; Indels 8; Gaps 4;  
 QY 13 ALDEETLWEM---MESHRRRIIVRCICPSRLTPYIRQAKVLCQDEEVLHSPRLTNSAM 68  
 DB 7 SLTEEDLTVKKDMLNRYVLCETKIIAERFHDLRKXKTLSEDTETI---SCR-TSSRX 63  
 QY 69 RAGHLDLTKRGKNGALAFLESK 93  
 DB 64 RAGKLDLYIQ-ENPRGLDTVESIR 87  
 RESULT 14  
 ID STYL\_HUMAN STANDARD; PRT; 685 AA.  
 AC Q13586;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Stromal interaction molecule 1 precursor.  
 GN STIM1 OR GOK.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Fetal liver, and Placenta;  
 RX MEDLINE=97079692; PubMed=8921403;  
 RA Parker N.U., Begley C.G., Smith P.J., Fox R.M.,  
 RT "Molecular cloning of a novel human gene (D11S4896E) at chromosomal  
 RT region 11p15.5.";  
 RL Genomics 37:253-256 (1996).  
 RN [2]  
 RP GLYCOSYLATION, PHOSPHORYLATION, SUBCELLULAR LOCATION, AND TISSUE  
 RP SPECIFICITY.  
 RX MEDLINE=20461006; PubMed=11004585;  
 RA Manji S.S., Parker N.J., Williams R.T., van Stekelenburg L.,  
 RA Pearson R.B., Dziadek M., Smith P.J.,  
 RT "STIM1: a novel phosphoprotein located at the cell surface.";  
 RL Biochim. Biophys. Acta 1481:147-155 (2000).  
 RN [3]  
 RP TISSUE SPECIFICITY, AND SUBUNIT.  
 RX MEDLINE=21356314; PubMed=11463338;  
 RA Williams R.T., Manji S.S.M., Parker N.J., Hancock M.S.,  
 RA Van Stekelenburg L., Eid U.-P., Senior P.V., Kazemwadi J.S.,  
 RA Shandala T., Saito R., Smith P.J., Dziadek M.A.,  
 RT "Identification and characterization of the STIM (stromal interaction  
 RT molecule) gene family: coding for a novel class of transmembrane  
 RT proteins.";  
 RL Biochem. J. 357:673-685 (2001).  
 RN [4]  
 RP DISEASE.  
 RX MEDLINE=98021968; PubMed=9377559;  
 RA Sabbioni S., Barbanti-Brodano G., Croce C.M., Negri M.,  
 RT "GOK: a gene at 11p15 involved in rhabdomyosarcoma and rhabdoid tumor  
 RT development.";  
 RL Cancer Res. 57:4493-4497 (1997).  
 CC -1- FUNCTION: Possibly adhesion molecule with a role in early  
 CC hematopoiesis by mediating attachment to stromal cells. Influences  
 CC the survival and/or proliferation of B cell precursors. Binding to  
 CC cells requires Mn(II) (By similarity).  
 CC -1- SUBUNIT: Oligomer with STIM2.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface  
 CC (Potential).  
 CC -1- TISSUE SPECIFICITY: Ubiquitously expressed in various human  
 CC primary cells and tumor cell lines.  
 CC -1- PTM: Glycosylated.  
 CC -1- PTM: Phosphorylated predominantly on Ser residues.



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 18, 2004, 04:30:32 ; Search time 85 Seconds

(without alignments)  
421.992 Million cell updates/sec

Title: US-10-032-159A-16

Perfect score: 720  
Sequence: 1 MGSLCRDRLTALDEBTLW.....SPFDGLAGTSRNRLILVTEX 139

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_protist:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriophage:\*  
17: sp\_archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	270.5	37.6	1159	11	Q8CIS0 mus musculu
2	270.5	37.6	1171	4	Q8TES3 homo sapien
3	215.5	29.9	535	13	Q8JFQ3 cyprinus ca
4	80	11.1	2114	10	Q9C6Y4 arabidopsi
5	77.5	10.8	2548	5	Q8I8S5 plasmodium
6	76.5	10.6	532	16	Q914C4 pseudomonas
7	75.5	10.5	887	12	Q9YV93 macropodid
8	74.5	10.3	394	2	Q9F5M0 pseudomonas
9	74	10.3	210	12	Q66577 equine he
10	74	10.3	311	12	Q9YV95 equine he
11	73.5	10.2	315	5	Q9NA19 fuscobacteri
12	73.5	10.2	382	16	Q8R6B3 caenorhabdi
13	73	10.1	265	16	Q97EL5 clostridium
14	73	10.1	447	16	Q8CV98 oceanobacill
15	72.5	10.1	295	5	Q76407 caenorhabdi
16	72.5	10.1	701	16	Q8CST1 staphylococ

17	72.5	10.1	845	5	Q9Y466	Q9Y466 drosophila
18	72.5	10.1	909	10	Q9M5A1	Q9M5A1 arabidopsis
19	72.5	10.1	1318	5	Q19733	Q19733 caenorhabdi
20	72	10.0	209	4	Q8N197	Q8N197 homo sapien
21	72	10.0	230	17	Q8TRF2	Q8TRF2 methanosarc
22	72	10.0	664	10	Q9S714	Q9S714 arabidopsis
23	72	10.0	1234	5	Q24690	Q24690 drosophila
24	71.5	9.9	660	10	Q9SNE4	Q9SNE4 arabidopsis
25	71.5	9.9	707	4	Q9NVK3	Q9NVK3 homo sapien
26	71.5	9.9	894	12	Q9YV92	Q9YV92 macropodid
27	71.5	9.9	956	4	Q9NV73	Q9NV73 homo sapien
28	71.5	9.9	956	4	Q8NEC0	Q8NEC0 homo sapien
29	71.5	9.9	1141	4	Q8CAE7	Q8CAE7 homo sapien
30	71.5	9.9	2099	4	Q9UNU8	Q9UNU8 homo sapien
31	71.5	9.9	2099	4	Q9Y2W9	Q9Y2W9 homo sapien
32	71	9.9	213	17	Q9YVZ2	Q9YVZ2 sulfobolus
33	71	9.9	265	5	Q61218	Q61218 caenorhabdi
34	71	9.9	310	16	Q8R888	Q8R888 thermococ
35	71	9.9	379	17	Q8T711	Q8T711 methanosarc
36	71	9.9	428	16	Q8ZF14	Q8ZF14 yersinia pe
37	70.5	9.8	263	11	Q60407	Q60407 cricetus cr
38	70.5	9.8	329	16	Q916H2	Q916H2 pseudomonas
39	70.5	9.8	735	10	Q91N85	Q91N85 arabidopsis
40	70.5	9.8	2215	10	Q8LH24	Q8LH24 oryza sativ
41	70	9.7	305	17	Q976C0	Q976C0 sulfobolus
42	70	9.7	308	10	Q8LFW3	Q8LFW3 helianthus
43	70	9.7	525	12	Q9MMJ0	Q9MMJ0 measles vir
44	70	9.7	599	16	Q8YH20	Q8YH20 bruceella me
45	70	9.7	741	12	Q9YIS6	Q9YIS6 xestia c-ni

## ALIGNMENTS

RESULT 1	Q8CIS0	PRELIMINARY;	PRT; 1159 AA.
ID	Q8CIS0		
AC	Q8CIS0		
DT	01-MAR-2003 (TREMBLrel. 23, Created)		
DE	01-MAR-2003 (TREMBLrel. 23, Last sequence update)		
DE	01-MAR-2003 (TREMBLrel. 23, Last annotation update)		
DE	Caspase recruitment domain family member 11.		
GN	CARD11.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OC	NCBI_TaxId=10090;		
RN	[1]		
RC	SEQUENCE FROM N.A.		
RC	TISSUE=Thymus;		
RA	MEDLINE=2244060; PubMed=12356734;		
RA	Pomerantz J.L., Denny E.M., Baltimore D.;		
RT	"CARD11 mediates factor-specific activation of NF-kappaB by the T cell		
RT	receptor complex."		
RU	EMBO J. 21:5184-5194(2002).		
DR	EMBL; AY135367; AAN10150.1; -		
SQ	SEQUENCE 1159 AA; 134039 MW; 8947A3CD5D0D81CD CRC64;		

Query Match	37.6%; Score 270.5; DB 11; Length 1159;
Best Local Similarity	48.6%; Pred. No. 1.3e-19;
Matches	54; Conservative 23; Mismatches 33; Indels 1; Gaps 1;
QY	15 DEETLWEMSHRRHIVRCICPSRLTPYLRQAKVLCOLDEEVLHSPRLTNSAPRAGHLL 74
DB	18 EEELALMDNVECNHRHSLRYINPAKLTPTLRQCKVLDQEDDEVLANAPMLPSKINRAGRL 77
QY	75 DLILTRGKNGAIAFLSLKLNHPDVTYITVIGLQPDVDSNFSGSSDDPGL 125
DB	78 DILHTGQGVVFLSLFLFYPPRLYKLVTKGKPTFRFSTIVVEEG-HEGL 127
RESULT 2	Q8TES3

```

ID 08TES3 PRELIMINARY; PRT; 1171 AA.
AC 08TES3;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE FLJ00120 protein (Fragment).
GN FLJ00120.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Ohara O., Nagase T., Kikuno R., Okumura K.;
RT "The nucleotide sequence of a long cDNA clone isolated from human
RT spleen."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK074049; BAB84875.1; -.
DR InterPro; IPR001315; CARD.
DR SMART; SM00228; PDZ; 1.
DR PROSITE; PS50209; CARD; 1.
FT NON_TER
SQ SEQUENCE 1171 AA; 134966 MW; FA567ABBC8A703FF CRC64;

Query Match 37.6%; Score 270.5; DB 4; Length 1171;
Best Local Similarity 48.6%; Pred. No. 1.3e-19;
Matches 54; Conservative 23; Mismatches 33; Indels 1; Gaps 1;

QY 15 DEEFLWEMESHRRIVRCICPSRLTPYLRQAKVLCQLEEEVYLSFRLTNSAMRAGHLL 74
DB 35 EEDALWENECNRMHSRYINPAKLTPIYARQCKVIDQEDDEVYANRPMTPSKINRAGRL 94

QY 75 DLKTRGKNGAIAFLSLKFNPDVYTLVTGLQPDVDFSNFGSESSDPDGL 125
DB 95 DILHTGGRGVVFLSLERYVELKLVTKGKPTFRSTIVESG-HBGL 144

RESULT 3
Q8FQ03 PRELIMINARY; PRT; 535 AA.
AC 08FQ03;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Caspase recruitment domain protein 9.
GN CARD9.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxId=7962;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Head kidney;
RA Kono T., Sakai M.;
RT "Molecular cloning of a novel caspase recruitment domain protein 9
RT (CARD9) in teleosts."
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB086061; BAC0527.1; -.
DR InterPro; IPR001315; CARD.
DR PROSITE; PS50209; CARD; 1.
SQ SEQUENCE 535 AA; 63375 MW; 9F3265D06C452D4D CRC64;

Query Match 29.9%; Score 215.5; DB 13; Length 535;
Best Local Similarity 38.2%; Pred. No. 3.3e-14;
Matches 52; Conservative 21; Mismatches 50; Indels 13; Gaps 2;

QY 15 DEEFLWEMESHRRIVRCICPSRLTPYLRQAKVLCQLEEEVYLSFRLTNSAMRAGHLL 74
DB 12 EDEECARLDEYMLIKITEPSRITPYLRQCKVLSSEDEQLYNDPSLVTRRKVGMILL 71

```

```

QY 75 DLKTRGKNGAIAFLSLKFNPDVYTLVTGLQPDVDFSNFGSESSDPDGL 121
DB 72 DILRTGKNGYEALFSLIEDLPDVYRKITGKEPARVSVLIDYAGCGTLQFLMSEVSR 131

QY 122 FDGLAGTSRNLRLVLT 137
DB 132 LQKLAQDERRRKLEVS 147

RESULT 4
Q9C6Y4 PRELIMINARY; PRT; 2114 AA.
AC 09C6Y4;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical 230.9 kDa protein.
GN T7023.25.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Euphorbia; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehlerr E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huzar L.,
RA Hunter J.W., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Lurco J.S., Malti R., Marzilli A.,
RA Miltseher J., Miranda M., Nguyen M., Nieman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utecherback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana."
RL Nature 408:816-820 (2000).
DR EMBL; AC074228; AAC50555.1; -.
DR InterPro; IPR000225; Armadillo.
DR InterPro; IPR000088; C2.
DR Pfam; PF00514; Armadillo_egg; 13.
DR Pfam; PF00168; C2; 1.
DR SMART; SM00185; ARM; 9.
DR SMART; SM00239; C2; 1.
DR Hypothetical protein.
SQ SEQUENCE 2114 AA; 230851 MW; 76E0C52A6B90A4 CRC64;

Query Match 11.1%; Score 80; DB 10; Length 2114;
Best Local Similarity 26.5%; Pred. No. 40;
Matches 40; Conservative 22; Mismatches 69; Indels 20; Gaps 5;

QY 4 LCRDASALTALDE--ETLWEMESHRR-----IVRCISRLTPYLRQAKVLCQ 52
DB 1206 LSPDSTETITVSELLESFSPETIRKTAISSKQIIGIHASRSRY-NARVVCEL 1264

QY 53 DEEFLWESHRRIVRCICPSRLTPYLRQAKVLCQLEEEVYLSFRLTNSAMRAGHLL 106
DB 1265 FSEHIRPSELAMKRL--SPLEWANTTLBESRYAALTALVXTLMGINPRDIITSLEGN 1322

QY 107 QPDVDFSNFGSESSDPDGLAGTSRNLRLVLT 137
DB 1323 PLDNIYKTLISLSSGSLSEKTSARICRFLT 1353

```





DR InterPro: IPR002938; MoxY\_FAD\_binding.  
 DR InterPro: IPR003042; Rng\_mnxygenase.  
 DR Pfam: PF01494; FAD\_binding\_3; 1.  
 DR Pfam: PF01360; Monooxygenase; 1.  
 DR PRINTS: PR00420; RINGMOXGNASE.  
 KM Oxidoreductase.  
 SQ SEQUENCE 394 AA; 44289 MW; F5E9354AE87A27A CRC64;

Query Match 10.3%; Score 74.5; DB 2; Length 394;  
 Best Local Similarity 24.8%; Pred. No. 19;  
 Matches 32; Conservative 17; Mismatches 49; Indels 31; Gaps 4;

QY 29 RIVRCICPSRLTYLROAKYLQDDEEVLHS-----PRLTNSAMRAQHLL----- 74  
 DB 42 RIRAGVLEQGMVLLREAGVGQMDREGVHDFELAFDGRLEIRIDRLTGTGKTYWYVG 101  
 QY 75 -----DLKTRGKNGAIAFLS-----LKFHNDVYTLVTGLQPDVDFNSFGSSSD 121  
 DB 102 QTEVTDLMFARAAGAPCFYEASEVELHKLKGENPHVFLHGGQMRIDCOQIAG----- 157  
 QY 122 FDGAGTSR 130  
 DB 158 CDGFGVSR 166

## RESULT 9

066677 PRELIMINARY; PRT; 210 AA.

AC 066677;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE ORF B10  
 OS Equine herpesvirus type 2 (strain 86/87) (EHV-2).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Gammaherpesvirinae.  
 CX NCBI\_TaxID=82831;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9302501; Pubmed=7783207;  
 RA Telford E.A., Watson M.S., Aird H.C., Perry J., Davison A.J.;  
 RT "The DNA sequence of equine herpesvirus 2.";  
 RL J. Mol. Biol. 249:520-526(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Telford E.A.R.;  
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U20824; AAC13865.1;  
 DR InterPro: IPR001315; CARD.  
 DR Pfam; PF00619; CARD; 1.  
 DR PROSITE; PS50209; CARD; 1.  
 SQ SEQUENCE 210 AA; 22943 MW; 136D639D8BC2D40 CRC64;

Query Match 10.3%; Score 74; DB 12; Length 210;  
 Best Local Similarity 28.0%; Pred. No. 9.6;  
 Matches 30; Conservative 14; Mismatches 27; Indels 36; Gaps 6;

QY 2 GELCRDSALTALDEETMEM-----MESHRRIRVRCICPSRLTPYLRQA 46  
 DB 10 GDPC-----VLTREBDIMVERLCBELRYLVLSHKSHK-----LDHLRAK 52  
 QY 47 KYLCQDDEEVLHSPRLTNSAMRAGHLLDLKTRGKNGAIAFLSLSK 93  
 DB 53 KILSRDAEYV--SSRAT--SRSRAGLLVDMCQDHP--GFQCLKESCK 95

RESULT 10  
 QYUN5 PRELIMINARY; PRT; 311 AA.  
 AC QYUN5;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

DE Hypothetical 32.6 kDa protein.  
 GN E10 OR CLAP.  
 OS Equine herpesvirus 2.  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Gammaherpesvirinae.  
 CX NCBI\_TaxID=133899;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99142601; Pubmed=9989495;  
 RA Willis T.G., Jadayel D.M., Du M.Q., Peng H., Perry A.R.,  
 RA Abdul-Rauf W., Price H., Karren L., Majekodunni O., Wlodarska I.,  
 RA Pan L., Crook T., Hamoudi R., Isaacson P., Dyer M.J.;  
 RT "Bcl10 is involved in t(1;14) (p22;q32) of MALT B cell lymphoma and  
 RT mutated in multiple tumor types.";  
 RL Cell 96:35-45(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Thome M., Martignon F., Hofmann K., Rubio V., Steiner V., Schneider P.,  
 RA Matman C., Tschopp J.;  
 RT "Equine herpesvirus-2 B10, but not its cellular homologue, activates  
 RT NF-kB transcription factor and c-Jun N-terminal kinase.";  
 RL J. Biol. Chem. 0:0-0(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99292766; Pubmed=10364242;  
 RA Sriivasula S.M., Ahmad M., Lin J.-H., Poyet J.-L.,  
 RA Fernandes-Alnemir T., Trichlis P.N., Alnemir E.S.;  
 RT "CLAP, a novel caspase recruitment domain-containing protein in the  
 RT tumor necrosis factor receptor pathway, regulates NF-kappaB activation  
 RT and apoptosis.";  
 RL J. Biol. Chem. 274:17946-17954(1999).  
 DR EMBL; AJ006410; CA87016.1;  
 DR EMBL; AF100340; AAD16430.1;  
 DR EMBL; AF134394; AAD39146.1;  
 DR InterPro: IPR001315; CARD.  
 DR Pfam; PF00619; CARD; 1.  
 DR PROSITE; PS50209; CARD; 1.  
 KM Hypothetical protein.  
 SQ SEQUENCE 311 AA; 32602 MW; 04EC68C735D723B CRC64;

Query Match 10.3%; Score 74; DB 12; Length 311;  
 Best Local Similarity 28.0%; Pred. No. 16;  
 Matches 30; Conservative 14; Mismatches 27; Indels 36; Gaps 6;

QY 2 GELCRDSALTALDEETMEM-----MESHRRIRVRCICPSRLTPYLRQA 46  
 DB 10 GDPC-----VLTREBDIMVERLCBELRYLVLSHKSHK-----LDHLRAK 52  
 QY 47 KYLCQDDEEVLHSPRLTNSAMRAGHLLDLKTRGKNGAIAFLSLSK 93  
 DB 53 KILSRDAEYV--SSRAT--SRSRAGLLVDMCQDHP--GFQCLKESCK 95

## RESULT 11

QYUN19 PRELIMINARY; PRT; 315 AA.

AC QYUN19;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Y38E10A.22 protein.  
 GN Y38E10A.22.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 CX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wallis J.M.;  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; Pubmed=9851916;

```

RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RL investigating biology.";
RL Science 282:2012-2018 (1998).
DR EMBL; AL110484; CAB54414.1; -.
DR WormPeP; Y38E10A.22; CER1603.
SQ SEQUENCE 315 AA; 36286 MW; 6f89223DFED00831 CRC64;

Query Match
Best Local Similarity 27.0%; Pred. No. 18;
Matches 31; Conservative 15; Mismatches 42; Indels 27; Gaps 5;

QY 23 MESHRIRIVRCICPSRL--TPYLROAKVLCQDDEEVLHSPRLTNSAMRA-----GHLD 75
DB 18 LKHNPEPQALADSPFLFAYVKKLKRQAE-----SPRLSLALSLSIYASHLPA 70

QY 76 LKTRGKNGALA--FLBSLKFHPDV-----YTLVTGLQPDVDFSNFSG 117
DB 71 YLKPAKNGLTATEKLIQSTCKDKNAVYAGRYIVSFLVTLTVKKNPVPFPADLQ 125

RESULT 12
Q8R6B3 PRELIMINARY; PRT; 382 AA.
AC Q8R6B3;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE (R)-2-Hydroxyglutaryl-CoA dehydratase beta-subunit (EC 4.2.1.-).
GN FN0208.
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteriia; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_Taxid=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=2186394; PubMed=11889109;
RA Kapural V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharya A., Bartman A., Gardner W., Greeklin G., Zhu L.,
RA Vasileva O., Chu L., Walunas T., Pusch G., Haselkorn R.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fongstein M., Kyriades N., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
RT nucleatum strain ATCC 25586."
RL J. Bacteriol. 184:2005-2016(2002).
DR EMBL; AE010533; AAL94414.1; -.
KW Lyase; Complete proteome.
SQ SEQUENCE 382 AA; 43923 MW; 93038D4296AE34CD CRC64;

Query Match
Best Local Similarity 26.9%; Pred. No. 23;
Matches 28; Conservative 15; Mismatches 44; Indels 17; Gaps 4;

QY 41 PYLROAKVLCQDDEEVLHSPRLTNSAMRAGHLLDLTKRNGKNGALF----- 89
DB 281 PMLALADQPARVDEPPLVDPPIK---RPKYVLDLVENNADGGLFMMNFNDTEKEY 337

QY 90 ESLK--FHNPDVYTLVTGL--QPDVDFSNFSGSSDFDGIAGTSR 130
DB 338 PSLKQAFBAKVPILRMGYDQGVDFGQVKTQLTFEFLVQLSR 381

RESULT 13
Q97EL5 PRELIMINARY; PRT; 265 AA.
AC Q97EL5;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hydroxyethylthiazole kinase Thim/Thik (FS1).
GN CAC3095.
OS Clostridium acetobutylicum.

```

```

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_Taxid=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breston G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hilti J., Wolf Y.I.,
RA Tatusov R.L., Sabatthe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum."
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AE007806; AAK81035.1; -.
DR InterPro; IPR004399; HMP-P_kinase.
DR TIGRFAMs; TIGR00097; HMP-P_kinase; 1.
KW Kinase; Complete proteome.
SQ SEQUENCE 265 AA; 28539 MW; 327DA3CEABED06EB CRC64;

Query Match
Best Local Similarity 23.8%; Pred. No. 16;
Matches 20; Conservative 23; Mismatches 29; Indels 12; Gaps 4;

QY 29 RIVRCICP--SKLTPYLROAKVLCQDDEEVLHSPRLTNS-----AMRAGHLD-L 76
DB 125 KLTCLPLDCLTLPITPPEAVLGGFKIKSEQMVPAAKIKSKINVLIKGSHVDNA 184

QY 77 LKTRGKNGALAFBSLKFHPDV 100
DB 185 TDLFYKNGYLEWFKLEKIDNPNT 208

RESULT 14
Q8CV98 PRELIMINARY; PRT; 447 AA.
AC Q8CV98;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE NADH oxidase.
GN OB0859.
OS Oceanobacillus ihyensis.
OC Bacteria; Firmicutes; Bacillales; Oceanobacillus.
OX NCBI_Taxid=182710;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HTE831 / DSM 14371 / JCM 11309;
RX MEDLINE=2220767; PubMed=12235376;
RA Takami H., Takaki Y., Uchiyama I.;
RT "Genome sequence of Oceanobacillus ihyensis isolated from the Ineya
RT ridge and its unexpected adaptive capabilities to extreme
RT environments."
RL Nucleic Acids Res. 30:3927-3935(2002).
DR EMBL; AP004596; BAC12815.1; -.
KW Complete proteome.
SQ SEQUENCE 447 AA; 48784 MW; 0388A6FBB4647773 CRC64;

Query Match
Best Local Similarity 26.8%; Pred. No. 32;
Matches 26; Conservative 15; Mismatches 50; Indels 6; Gaps 1;

QY 21 EMESHRIIVRCICPSRLTPYLROAKVLCQDDEEVLHSPRLTNSAMRAGHLLDLKTR 80
DB 163 EMASLCHRGFSCSLVD-----RSEHVLKRIKEMAHIDELQKGIALVYNDGLKSF 216

QY 81 GKNGALAFBSLKFHPDVYTLVTGLQPDVDFSNFSG 117
DB 217 SDNGTLLSSDKTIQADMTVAIGIKPNTETLAIDAG 253

RESULT 15
Q76407

```

```

ID 076407 PRELIMINARY; PRT; 295 AA.
AC 076407;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical protein T10B5.3.
GN T10B5.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Waterston R.;
RT "Genome sequence of the nematode C. elegans: a platform for
   investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Geisel C., Bradshaw H., O'Brien D.;
RT "The sequence of C. elegans cosmid T10B5.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RX Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF067947; AAK68397.2; -.
DR Wormpep; T10B5.3; CE30832.
KM Hypothetical protein.
SQ SEQUENCE 295 AA; 32436 MW; 637C20FC0055AA02 CRC64;

```

Query March 10.1%; Score 72.5; DB 5; Length 295;

Best Local Similarity 24.3%; Pred. No. 21;

Matches 35; Conservative 20; Mismatches 48; Indels 41; Gaps 6;

```

QY 8 DSALTALDEFTLWEMESHRRHIVRCIPSRLLTPYLR-----QAKYLCQLDREEV 57
Db 100 DSVIRNINSEII-----HATLRDLQPMQVPLPKITSRUKTRNADIRPTIRMAQI 151
QY 58 ---LHSPRLTN-----SAMRAGHLIDLKTRGKGAIAFLESLEKRNPDVY 100
Db 152 AFSIHMEFYLSSLPTMEKEIGGLIGMLRSRVGHHRDILALHGKISTIDLIKRRNTN----- 207
QY 101 TLVTGLQPDVDFSN-FSGESSDFD 123
Db 208 -VVIWQQLVFNNDLDSSEDFD 230

```

Search completed: February 18, 2004, 04:46:49  
 Job time : 91 secs



CURRENT FILING DATE: 1998-04-27  
NUMBER OF SEQ ID NOS: 38  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 38  
LENGTH: 233  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-069-023-38

Query Match 10.8%; Score 78; DB 4; Length 233;  
Best Local Similarity 35.3%; Pred. No. 0.09;  
Matches 30; Conservative 14; Mismatches 33; Indels 8; Gaps 4;

QY 13 ALDEETMEM---MESHRRIVRCIPSRITPYLRQAKVLCQDEEVLHSPRLTNSAM 68  
DB 7 SLTEEDLEVEKDALENIRVLCERKIIAEHFHDLRAKILSREDTEI--SCR-TSSRX 63  
QY 69 RAGHLDLKTRGKGAIAFLSLK 93  
DB 64 RAGKLDVLC-ENPRGLDTEVSR 87

RESULT 3  
US-09-252-991A-19231  
Sequence 19231, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252, 991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 19231  
LENGTH: 621  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-19231

Query Match 10.6%; Score 76.5; DB 4; Length 621;  
Best Local Similarity 26.7%; Pred. No. 0.63;  
Matches 40; Conservative 18; Mismatches 67; Indels 25; Gaps 6;

QY 4 LCRASALTA-----LDEETLWEMESHRRIVRC-----ICPSRLTPYLRLQ 45  
DB 286 LQORRLATAPASEVDSIDEQ-IGETLRCROTFFPCALVSGVDVSLGSLYDLPOLR 344  
QY 46 AKVLCQDEEVLHSPRLTNSAMRAGHLLDKTR-GRNGAIAFLSLKFNPDVYTLVT 104  
DB 345 FHLCEGEBESLPHPRLOFELRQEAFFPLLRKAVNGFQATMSILMTQ---RLAD 400  
QY 105 GLQPDVDFSNFSGESSDPDGLAGTSRNLRL 134  
DB 401 GIGEGYHCVLIGEGAD-ELFWGYPRLHLEL 429

RESULT 4  
US-09-618-425-9  
Sequence 9, Application US/09618425  
Patent No. 5268278  
APPLICANT: CANOSI, UMBERTO, DE FAZIO, GABRIELE, VILLA, STEFANO,  
DONINI, SILVIA  
TITLE OF INVENTION: GENETIC EXPRESSION OF SOMATOSTATIN AS  
HYBRID POLYPEPTIDE  
NUMBER OF SEQUENCES: 4  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/193,202  
FILING DATE: 09-MAY-1988  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 717,444  
FILING DATE: 29-MAR-1985  
SEQ ID NO:3;  
LENGTH: 326  
5268278-3

Query Match 10.6%; Score 76; DB 6; Length 326;  
Best Local Similarity 27.0%; Pred. No. 0.27;  
Matches 37; Conservative 22; Mismatches 58; Indels 20; Gaps 7;

QY 9 SALTLADEB---TLWEMESHRRIVRCIPSRITPYLRQAKVLCQD-EEEVLHSPRL 63  
DB 195 ASLPAPDEERKQRLTARLDELRLQQLTEAAP---LPVVSVPXMRCECQSDSEFGVRL 251  
QY 64 TNSAMRAGHLLDLKTRGN---GALFSLKFNPDVYTLVTGLQPDVDFSNF--SG 117  
DB 252 LQKAIKRAEIGFQVVSRRFSLPCPSPLAAYVLKKSDFSPYWF--FMQDNDFTLFGASP 308  
QY 118 ESSDPDGLAGTSRNLRL 134  
DB 309 ESSKD--ATSRQIEI 322

RESULT 5  
US-09-328-352-7785  
Sequence 7785, Application US/09328352  
Patent No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
FILE REFERENCE: GTC99-03PA  
CURRENT APPLICATION NUMBER: US/09/328,352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 7785  
LENGTH: 333  
TYPE: PRT  
ORGANISM: Acinetobacter baumannii  
US-09-328-352-7785

Query Match 10.3%; Score 74; DB 4; Length 333;  
Best Local Similarity 27.9%; Pred. No. 0.51;  
Matches 39; Conservative 14; Mismatches 41; Indels 46; Gaps 8;

QY 11 LTLADEETLWEMESHRRIVRCI-----CPSRLTPY-----LRQAKVLCQDEE 56  
DB 135 ITSDVDRDLHDGGAH---FVKCIEIRNRCPDTLIEIIVDPGRRLTALSTLSISPD 191  
QY 57 VHS-----PRLTNSAMRAG---HLLDLK-----TRGNG-----ALAF 88  
DB 192 VFNHNTETVRL-YKAMRPGSDVGHSLNLLKPKKVCPIKIKRGGLWGLGEIARVIL 250  
QY 89 LESLKFHNPDPVYTLVTGLQ 108  
DB 251 LNDLKDHVDVLTITGQYLP 270

RESULT 6  
US-09-618-425-9  
Sequence 9, Application US/09618425  
Patent No. 6475744  
GENERAL INFORMATION:  
APPLICANT: Reppert, Steven M.  
APPLICANT: Weaver, David R.  
APPLICANT: Zylka, Mark  
APPLICANT: Jin, Xiaowei  
APPLICANT: Kume, Kazuhiko  
APPLICANT: Siram, Satyanarayanan  
APPLICANT: Sheatman, Lauren  
TITLE OF INVENTION: METHODS FOR IDENTIFYING COMPOUNDS WHICH  
MODULATE CIRCADIAN RHYTHM  
FILE REFERENCE: 00786-428001

; CURRENT APPLICATION NUMBER: US/09/618,425  
 ; CURRENT FILING DATE: 2000-07-18  
 ; PRIOR APPLICATION NUMBER: 60/203,005  
 ; PRIOR FILING DATE: 2000-05-10  
 ; PRIOR APPLICATION NUMBER: 60/145,363  
 ; PRIOR FILING DATE: 1999-07-22  
 ; NUMBER OF SEQ ID NOS: 13  
 ; SOFTWARE: FASTSEQ for Windows Version 4.0  
 ; SEQ ID NO 9  
 ; LENGTH: 1113  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 ; US-09-618-425-9

Query Match 10.2%; Score 73.5; DB 4; Length 1113;  
 Best Local Similarity 25.2%; Pred. No. 3.8;  
 Matches 26; Conservative 19; Mismatches 33; Indels 25; Gaps 4;

QY 6 RDSALTLADEETLWEMESHRRIVRCICPSRLTPYRQAKVCO-----LDEEVL 58  
 DB 1003 QREDELPGAABESITWRME---RTPECVMTYQVPERGREEVKODLEKLSMEQOQPL 1058  
 QY 59 HSPRL-----TNSAMRAGHLIDL--KTRKNGAIA 87  
 DB 1059 FSPQREELAKVRWISHHTAPQESHTQSCVACEDRGSVDTA 1101

RESULT 7  
 ; US-09-134-001C-4908  
 ; Sequence 4908, Application US/09134001C  
 ; Patent No. 6380370  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn Doucette-Stamm et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
 ; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: GTC-007  
 ; CURRENT APPLICATION NUMBER: US/09/134,001C  
 ; CURRENT FILING DATE: 1998-08-13  
 ; PRIOR APPLICATION NUMBER: US 60/064,964  
 ; PRIOR FILING DATE: 1997-11-08  
 ; PRIOR APPLICATION NUMBER: US 60/055,779  
 ; PRIOR FILING DATE: 1997-08-14  
 ; NUMBER OF SEQ ID NOS: 5674  
 ; SEQ ID NO 4908  
 ; LENGTH: 706  
 ; TYPE: PRT  
 ; ORGANISM: Staphylococcus epidermidis  
 ; US-09-134-001C-4908

Query Match 10.1%; Score 72.5; DB 4; Length 706;  
 Best Local Similarity 25.4%; Pred. No. 2.5;  
 Matches 36; Conservative 18; Mismatches 47; Indels 41; Gaps 6;

QY 6 RDSALTLADEETLWEMESH-----RHRIVRCICPSRLTPYRQAKV 46  
 DB 271 QREDELTLKKEVVGHFDEEDPENETVKEVVAINDLIGEVRRLLIADKXIPDGRKV 330  
 QY 47 KVLQDLDEEVLHSPRLTNSAMRAGHLIDLKTRKNGAIAFLSKFHPNDVYTLVTL 106  
 DB 331 DEIRPL-ESVGLLPRAGSG-----LFTNGQTALSVLTGALGD---YQLIDGL 377  
 QY 107 QPDVD-----FSNFS-GES 119  
 DB 378 GPEVEKRFMEHYNFPNFSVGET 399

RESULT 8  
 ; US-09-252-991A-27778  
 ; Sequence 27778, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 107196,136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 27778  
 ; LENGTH: 614  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas aeruginosa  
 ; US-09-252-991A-27778

Query Match 9.8%; Score 70.5; DB 4; Length 614;  
 Best Local Similarity 25.2%; Pred. No. 3.7;  
 Matches 31; Conservative 17; Mismatches 44; Indels 31; Gaps 5;

QY 46 AKVLCQJDEEE---VLHSP-----RLTNSAMRAGHLIDLKTRG--- 81  
 DB 368 AKPLQGLEEETSDLAIVPATRTLTFTGKRLVLESLTGDVLRTPILGLSPGVAV 427  
 QY 82 -KNGAIAFLSKL-----TH-NDVYTLVTLGLOPDVDFSNFSGSSDPGLACTSRNL 134  
 DB 428 LKNGVAIVDERNTLTIFVDPQTAELSTKSLAEPLPLNIGKNGKFGIAWDPPQORL 487

QY 135 LVT 137  
 DB 488 LLS 490

RESULT 9  
 ; US-08-652-971-4  
 ; Sequence 4, Application US/08652971  
 ; Patent No. 5814507  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cheng, Jili  
 ; APPLICANT: Laeky, Laurence A.  
 ; TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE  
 ; TITLE OF INVENTION: PHOSPHATASE, PTP LAMEDA  
 ; NUMBER OF SEQUENCES: 10  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Genentech, Inc.  
 ; STREET: 460 Point San Bruno Blvd.  
 ; CITY: South San Francisco  
 ; STATE: California  
 ; COUNTRY: United States  
 ; ZIP: 94080  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/652,971  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Dreger, Ginger R.  
 ; REGISTRATION NUMBER: 33,055  
 ; REFERENCE/DOCKET NUMBER: P1033  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 225-3216  
 ; TELEFAX: (415) 952-9881  
 ; TELEX: 910 371-7168  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1452 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein

US-08-652-971-4

Query Match 9.7%; Score 69.5; DB 2; Length 1452;  
Best Local Similarity 26.9%; Pred. No. 19;  
Matches 32; Conservative 14; Mismatches 50; Indels 23; Gaps 6;

QY 16 EETLWEMESH-RHRIYRCICPSRLTPYLROAKYLCOLEDEEVLAHSPTLNSMRAGHLL 74  
DB 431 EEVSWDTDNHSHPOHTI-----TNLSPTNVSVKLLIMNPKGKESQELT--VQTDDELIP 482  
QY 75 DLKTRKNGAIAFLBS--LKFNPD-----VYTLVTGLQPDVDFSNFSGESS 120  
DB 483 GAVPTESIQGS-AFEKXIFLOWREPOTYGVITLYETTYKAVVSFPEDIDLSNOSGRVS 540

RESULT 10

US-08-991-258A-4  
Sequence 4, Application US/08991258A  
Patent No. 5928887

GENERAL INFORMATION:

APPLICANT: Cheng, Jili

APPLICANT: Lasky, Laurence A.

TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON &amp; HERBERT, LLP

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: United States

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/991,258A

FILING DATE: 17-DEC-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/652,971

FILING DATE: 24-MAY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Dreger, Walter H.

REGISTRATION NUMBER: 24,190

REFERENCE/DOCKET NUMBER: A-63478-3/WHD/MTX

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 781-1989

TELEFAX: (415) 398-3249

TELEX:

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 1452 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-991-258A-4

Query Match 9.7%; Score 69.5; DB 2; Length 1452;

Best Local Similarity 26.9%; Pred. No. 19;

Matches 32; Conservative 14; Mismatches 50; Indels 23; Gaps 6;

QY 16 EETLWEMESH-RHRIYRCICPSRLTPYLROAKYLCOLEDEEVLAHSPTLNSMRAGHLL 74  
DB 431 EEVSWDTDNHSHPOHTI-----TNLSPTNVSVKLLIMNPKGKESQELT--VQTDDELIP 482  
QY 75 DLKTRKNGAIAFLBS--LKFNPD-----VYTLVTGLQPDVDFSNFSGESS 120  
DB 483 GAVPTESIQGS-AFEKXIFLOWREPOTYGVITLYETTYKAVVSFPEDIDLSNOSGRVS 540

RESULT 11

US-08-769-399-4  
Sequence 4, Application US/08769399  
Patent No. 5976852

GENERAL INFORMATION:

APPLICANT: Cheng, Jili

APPLICANT: Lasky, Laurence A.

TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd.

CITY: South San Francisco

STATE: California

COUNTRY: United States

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/769,399

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Dreger, Ginger R.

REGISTRATION NUMBER: 33,055

REFERENCE/DOCKET NUMBER: P1033

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 225-3216

TELEFAX: (415) 952-9881

TELEX: 910 371-7168

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 1452 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-769-399-4

Query Match 9.7%; Score 69.5; DB 2; Length 1452;

Best Local Similarity 26.9%; Pred. No. 19;

Matches 32; Conservative 14; Mismatches 50; Indels 23; Gaps 6;

QY 16 EETLWEMESH-RHRIYRCICPSRLTPYLROAKYLCOLEDEEVLAHSPTLNSMRAGHLL 74  
DB 431 EEVSWDTDNHSHPOHTI-----TNLSPTNVSVKLLIMNPKGKESQELT--VQTDDELIP 482  
QY 75 DLKTRKNGAIAFLBS--LKFNPD-----VYTLVTGLQPDVDFSNFSGESS 120  
DB 483 GAVPTESIQGS-AFEKXIFLOWREPOTYGVITLYETTYKAVVSFPEDIDLSNOSGRVS 540

RESULT 12  
US-08-991-953A-4  
Sequence 4, Application US/08991953A  
Patent No. 6083748

GENERAL INFORMATION:  
APPLICANT: Cheng, Jili  
APPLICANT: Lasky, Laurence A.  
TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT, LLP  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: United States



ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/991,953A  
FILING DATE: 16-DEC-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/552,971  
FILING DATE: 24-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Dreger, Walter H.  
REGISTRATION NUMBER: 24,190  
REFERENCE/DOCKET NUMBER: A-63478-3/WHD/MTK  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1452 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
US-08-991-953A-4

Query Match 9.7%; Score 69.5; DB 3; Length 1452;  
Best Local Similarity 26.9%; Pred. No. 19;  
Matches 32; Conservative 14; Mismatches 50; Indels 23; Gaps 6;

QY 16 ESTLWEMESH-RRIYVRCIPSLTPYLRQAKVLCQDEEEVLSHSPRLTNSAMRAGHL 74  
DB 431 EVSWDTNDSHPQHT-----TMSPTYNVSVKILMNPGRKRSQELT--VQTDLELP 482

QY 75 DLKTRGNKGAIAPLES--LKFHNP-----YTLVTGLQPDVPSNFGSS 120  
DB 483 GAVPTESIQGS-APEKFIQWREPTQYGVITLYEITYKAVSFDPEIDLSNGRVS 540

RESULT 13  
US-09-198-452A-196  
; Sequence 196, Application US/09198452A  
; Patent No. 6559294  
; GENERAL INFORMATION:  
; APPLICANT: Griffaia, R.  
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments  
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention  
; FILE REFERENCE: 9710-003-999  
; CURRENT APPLICATION NUMBER: US/09/198,452A  
; CURRENT FILING DATE: 1998-11-24  
; NUMBER OF SEQ ID NOS: 6849  
; SEQ ID NO 196  
; LENGTH: 203  
; TYPE: PRT  
; ORGANISM: Chlamydia pneumoniae  
US-09-198-452A-196

Query Match 9.6%; Score 69; DB 4; Length 203;  
Best Local Similarity 21.1%; Pred. No. 1.1; Indels 20; Gaps 3;  
Matches 24; Conservative 23; Mismatches 47;

QY 28 HRIYACIPSLTPYLRQAKVLCQDEEEVLSHSPRLTNSA-----NEAGHLDLK 78  
DB 4 HXILDG-----KXKVFIGNVASEETPSQILHAGSEVADAFNQGADCHPAANQILBAK 59

QY 79 TRGNKGAIAFLAESLKFHNPDYTLVTGLQPDVPSNFGSSDPDGLAGTSRNL 132  
DB 60 -----ICLDVYHTNHYSVTFPCVDNYPNILRFYVSSKNENMGJSLNPDNY 106

RESULT 14  
US-09-198-122-2  
; Sequence 2, Application US/09198122  
; Patent No. 6180380  
; GENERAL INFORMATION:  
; APPLICANT: Streibardt, Klaus; Rubsamen-Waigmann, Helga;  
; APPLICANT: Holtzrich, Uwe  
; TITLE OF INVENTION: CLONING OF A MEMBER OF THE SERINE-  
; TITLE OF INVENTION: THREONINE-KINASE FAMILY  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SPRUNG HORN KRAMER & WOODS  
; STREET: 660 White Plains Road  
; CITY: Tarrytown  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10591-5144

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB  
MEDIUM TYPE: Storage  
COMPUTER: NEC Powermate SX-20  
OPERATING SYSTEM: DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/198,122  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/601,014  
FILING DATE: 23-FEB-1996  
APPLICATION NUMBER: PCT/EP94/02863  
FILING DATE: 30-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 43292177  
FILING DATE: 30-AUG-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Kurt G. Briscoe  
REGISTRATION NUMBER: 33,141  
REFERENCE/DOCKET NUMBER: Bayer 9516-KGB  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (914) 332-1700  
TELEFAX: (914) 332-1844  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 603 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
US-09-198-122-2

Query Match 9.6%; Score 69; DB 3; Length 603;  
Best Local Similarity 24.5%; Pred. No. 5.6; Indels 52; Gaps 6;  
Matches 34; Conservative 15; Mismatches 38;

QY 3 ELGRDSALTADEELTWMESHRIYRCIPSLTPYLRQAKVLCQDEEEVLSHSP 62  
DB 131 ELGRDS-----LLEPHKR--KALTEPARYYLRQIVGQ-----YLHNR 171

QY 63 LTNAMRAGHL---DLKTRGNKGAIAPLES-----SLKFNPD 98  
DB 172 VHRDKLGNLFNEDLVKIGDFGLATRYVDGERKTLGPNYIAPBVLSKHSP 231

QY 99 V-----YTLVTGLQPD 108  
DB 232 VDWVSGICINYYTLVGRKP 250

RESULT 15  
US-09-139-802-201  
; Sequence 201, Application US/09139802

```

; Patent No. 6180084
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: NGR Receptor and Methods of Identifying Tumor Homing
; TITLE OF INVENTION: Molecules That Home to Angiogenic Vasculature Using
; TITLE OF INVENTION: Same
; FILE REFERENCE: P-IJ 3203
; CURRENT APPLICATION NUMBER: US/09/139, 802
; CURRENT FILING DATE: 1998-08-25
; EARLIER APPLICATION NUMBER: 08/926, 914
; EARLIER FILING DATE: 1997-09-10
; EARLIER APPLICATION NUMBER: 08/710, 067
; EARLIER FILING DATE: 1996-09-10
; NUMBER OF SEQ ID NOS: 226
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 201
; LENGTH: 967
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-139-802-201

```

```

Query Match          9.6%; Score 69; DB 3; Length 967;
Best Local Similarity 25.0%; Pred. No. 12;
Matches 40; Conservative 25; Mismatches 51; Indels 44; Gaps 11;

```

```

QY 9 SALFALDEETIMEMESHRRIRVICPS---RLTPYL-----ROAKYLC- 50
DB 60 ASATTLDDSKAW---NRYRLPNTLKPDSYQVTLRPYLTPNDRLGLYFKGSSVTFPTCK 114
QY 51 QLDEEEVLAHPRLTNSAMRAGHLIDLKTRGKGA-IAFLESLEKFNPDVY-----TL 102
DB 115 EATVILIIHSKGL-NYTLISQHRVYL---KGVGSSQEPDIDKTELVETRYLVVHLKSL 170
QY 103 VTGLQPDVDFSNFSGSSDPDGLAGTSR-----NRLIV 136
DB 171 VKDSQYEMD-SEFEGELA--DLAGFYRSEYMEGNVRKV 207

```

```

Search completed: February 18, 2004, 04:48:20
Job time : 31 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 18, 2004, 04:46:58 ; Search time 76 Seconds  
(without alignments)  
382.949 Million cell updates/sec

Title: US-10-032-159a-16

Perfect score: 720  
Sequence: 1 MGELCRDSALTPALDEETLM.....SDPDLGAGTNRNLRLVTPX 139

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10C\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	719	99.9	139	US-10-032-159a-16	Sequence 16, Appl
2	714	99.2	138	US-09-767-215-5	Sequence 5, Appl1
3	617.5	85.8	418	US-10-104-047-2285	Sequence 2285, Ap
4	617.5	85.8	1004	US-09-767-215-2	Sequence 2, Appl1
5	484	67.2	92	US-10-032-159a-18	Sequence 16, Appl
6	270.5	37.6	1147	US-09-798-412-11	Sequence 11, Appl1
7	270.5	37.6	1147	US-10-325-917-11	Sequence 8, Appl1
8	270.5	37.6	1247	US-10-032-159a-8	Sequence 10, Appl
9	257	35.7	92	US-10-032-159a-10	Sequence 8, Appl1
10	246	34.2	1032	US-09-798-412-8	Sequence 10, Appl
11	246	34.2	1032	US-10-325-917-8	Sequence 8, Appl1
12	223.5	31.0	536	US-09-798-412-2	Sequence 2, Appl1
13	223.5	31.0	536	US-10-325-917-2	Sequence 2, Appl1
14	215.5	29.9	366	US-10-032-159a-2	Sequence 20, Appl
15	215.5	29.9	366	US-10-032-159a-20	Sequence 20, Appl

16	215.5	29.9	536	11	US-09-798-412-5	Sequence 5, Appl1
17	215.5	29.9	536	15	US-10-325-917-5	Sequence 5, Appl1
18	158	21.9	72	14	US-10-032-159a-4	Sequence 4, Appl1
19	113	15.7	48	12	US-10-029-386-34276	Sequence 34276, A
20	80	11.1	966	12	US-10-099-322-72	Sequence 72, Appl
21	80	11.1	966	12	US-10-044-564-72	Sequence 58, Appl
22	75.5	10.5	685	11	US-09-877-843-58	Sequence 12060, A
23	74.5	10.3	281	15	US-10-156-761-12060	Sequence 12989, A
24	74.5	10.3	329	12	US-10-369-493-12989	Sequence 9, Appl1
25	73.5	10.2	1113	15	US-10-245-175-9	Sequence 6948, Ap
26	72.5	10.1	1318	12	US-10-369-493-6948	Sequence 44, Appl
27	71.5	9.9	889	11	US-09-567B-44	Sequence 6408, Ap
28	70.5	9.8	356	12	US-10-369-493-6408	Sequence 196, App
29	69.5	9.7	431	14	US-10-078-929-200	Sequence 4617, Ap
30	69	9.6	203	12	US-10-289-762-196	Sequence 186, App
31	69	9.6	395	12	US-10-108-260A-4617	Sequence 71, Appl
32	69	9.6	603	15	US-10-171-311-186	Sequence 14, Appl
33	69	9.6	966	12	US-10-099-322-71	Sequence 75, Appl
34	69	9.6	966	12	US-10-044-564-71	Sequence 74, Appl
35	69	9.6	967	12	US-10-099-322-14	Sequence 75, Appl
36	69	9.6	967	12	US-10-099-322-74	Sequence 32, Appl
37	69	9.6	967	12	US-10-099-322-75	Sequence 201, Appl
38	69	9.6	967	12	US-10-044-564-14	Sequence 122, App
39	69	9.6	967	12	US-10-044-564-74	Sequence 73, Appl
40	69	9.6	967	12	US-10-044-564-75	
41	69	9.6	967	15	US-10-205-823-32	
42	69	9.6	967	15	US-10-264-374-201	
43	69	9.6	967	15	US-10-281-904-2	
44	69	9.6	969	10	US-09-981-353-122	
45	69	9.6	974	12	US-10-099-322-73	

#### ALIGNMENTS

RESULT 1  
US-10-032-159a-16  
; Sequence 16, Application US/10032159A  
; Publication No. US20020164703A1  
; GENERAL INFORMATION:  
; APPLICANT: Pawlowski, Krzysztof  
; APPLICANT: Reed, John C.  
; TITLE OF INVENTION: CARD-DOMAIN CONTAINING POLYPEPTIDES,  
; TITLE OF INVENTION: ENCODING NUCLEIC ACIDS, AND METHODS OF USE  
; FILE REFERENCE: P-11 5100  
; CURRENT APPLICATION NUMBER: US/10/032.159A  
; PRIOR FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: US 60/257,457  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 139  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: 139  
; OTHER INFORMATION: Xaa = Any Amino Acid  
US-10-032-159a-16  
Query Match 99.9%; Score 719; DB 14; Length 139;  
Best Local Similarity 100.0%; Pred. No. 6.1e-77;  
Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 61 PRLTNSMRAGHLLDLKTRGKNGAIAFLSLKFNHNDVTVTVGLOPDVDFSNFSGESS 120  
61 MGELCRDSALTPALDEETLMEMESHRRIVRCICPSRLTPYVROAKVLCOLBEETVHS 60  
1 MGELCRDSALTPALDEETLMEMESHRRIVRCICPSRLTPYVROAKVLCOLBEETVHS 60  
61 PRLTNSMRAGHLLDLKTRGKNGAIAFLSLKFNHNDVTVTVGLOPDVDFSNFSGESS 120  
61 PRLTNSMRAGHLLDLKTRGKNGAIAFLSLKFNHNDVTVTVGLOPDVDFSNFSGESS 120

Qy 121 DFDGAGTSRNLRLVTP 138  
Db 121 DFDGAGTSRNLRLVTP 138

## RESULT 2

US-09-767-215-5  
; Sequence 5, Application US/09767215  
; Patent No. US20020081636A1  
; GENERAL INFORMATION:  
; APPLICANT: Bertin, John  
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED  
; FILE REFERENCE: 07334-142001  
; CURRENT APPLICATION NUMBER: US/09/767,215  
; CURRENT FILING DATE: 2001-01-22  
; PRIOR FILING DATE: 2000-02-09  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 1138  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-767-215-5

Query Match 99.2%; Score 714; DB 9; Length 1138;  
Best Local Similarity 100.0%; Pred. No. 4, 5e-75;  
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GELCRDSALTLADDETLWEMESHRRIRIVRCICPSRLTPYLROAKYLCQDDEEVVHSP 61  
Db 1 GELCRDSALTLADDETLWEMESHRRIRIVRCICPSRLTPYLROAKYLCQDDEEVVHSP 60  
Qy 62 RLTSNAPRAGHLLDLKTRGNKGAIAFLSKFNPDPVYTLVTGLQPDVDFSNFSGSSD 121  
Db 61 RLTSNAPRAGHLLDLKTRGNKGAIAFLSKFNPDPVYTLVTGLQPDVDFSNFSGSSD 120  
Qy 122 FDGLAGTSRNLRLVTP 138  
Db 121 FDGLAGTSRNLRLVTP 137

## RESULT 3

US-10-104-047-2285  
; Sequence 2285, Application US/10104047  
; Publication No. US20030236392A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: NO. US20030236392A1 full length cDNA  
; FILE REFERENCE: H1-A0105  
; CURRENT APPLICATION NUMBER: US/10/104,047  
; CURRENT FILING DATE: 2002-03-25  
; PRIOR APPLICATION NUMBER:  
; PRIOR FILING DATE:  
; NUMBER OF SEQ ID NOS: 4096  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 2285  
; LENGTH: 418  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-104-047-2285

Query Match 85.8%; Score 617.5; DB 12; Length 418;  
Best Local Similarity 90.4%; Pred. No. 2, 9e-64;  
Matches 123; Conservative 4; Mismatches 6; Indels 3; Gaps 2;

Qy 1 MGELCRDSALTLADDETLWEMESHRRIRIVRCICPSRLTPYLROAKYLCQDDEEVVHSP 60  
Db 1 MGELCRDSALTLADDETLWEMESHRRIRIVRCICPSRLTPYLROAKYLCQDDEEVVHSP 60  
Qy 61 RLTSNAPRAGHLLDLKTRGNKGAIAFLSKFNPDPVYTLVTGLQPDVDFSNFSG--E 118

Db 61 RLTSNAPRAGHLLDLKTRGNKGAIAFLSKFNPDPVYTLVTGLQPDVDFSNFSGME 120  
Qy 119 SSDF-DGLAGTSRNLRL 133  
Db 121 TSKLTECLAGATGSLQ 136

## RESULT 4

US-09-767-215-2  
; Sequence 2, Application US/09767215  
; Patent No. US20020081636A1  
; GENERAL INFORMATION:  
; APPLICANT: Bertin, John  
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED  
; FILE REFERENCE: 07334-142001  
; CURRENT APPLICATION NUMBER: US/09/767,215  
; CURRENT FILING DATE: 2001-01-22  
; PRIOR FILING DATE: 2000-02-09  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 1004  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-767-215-2

Query Match 85.8%; Score 617.5; DB 9; Length 1004;  
Best Local Similarity 90.4%; Pred. No. 9, 9e-64;  
Matches 123; Conservative 4; Mismatches 6; Indels 3; Gaps 2;

Qy 1 MGELCRDSALTLADDETLWEMESHRRIRIVRCICPSRLTPYLROAKYLCQDDEEVVHSP 60  
Db 1 MGELCRDSALTLADDETLWEMESHRRIRIVRCICPSRLTPYLROAKYLCQDDEEVVHSP 60  
Qy 61 RLTSNAPRAGHLLDLKTRGNKGAIAFLSKFNPDPVYTLVTGLQPDVDFSNFSG--E 118  
Db 61 RLTSNAPRAGHLLDLKTRGNKGAIAFLSKFNPDPVYTLVTGLQPDVDFSNFSGME 120  
Qy 119 SSDF-DGLAGTSRNLRL 133  
Db 121 TSKLTECLAGATGSLQ 136

## RESULT 5

US-10-032-159a-18  
; Sequence 18, Application US/10032159A  
; Publication No. US20020164703A1  
; GENERAL INFORMATION:  
; APPLICANT: Pawlowski, Krzysztof  
; APPLICANT: Reed, John C.  
; TITLE OF INVENTION: CARD-DOMAIN CONTAINING POLYPEPTIDES,  
; FILE REFERENCE: P-LJ 5100  
; CURRENT APPLICATION NUMBER: US/10/032,159A  
; CURRENT FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: US 60/257,457  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 18  
; LENGTH: 92  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-032-159a-18

Query Match 67.2%; Score 484; DB 14; Length 92;  
Best Local Similarity 100.0%; Pred. No. 2, 2e-49;  
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	QY	QY	QY
16	16	16	16
1	1	1	1
76	76	76	76
61	61	61	61

RESULT 6  
US-09-798-412-11

```

1 GENERAL INFORMATION:
2 APPLICANT: Bertin, John
3 TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
4 TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
5 FILE REFERENCE: 07334-327001
6 CURRENT APPLICATION NUMBER: US/09/798,412
7 CURRENT FILING DATE: 2001-03-02
8 PRIOR APPLICATION NUMBER: US 09/728,260
9 PRIOR FILING DATE: 2000-12-01
10 PRIOR APPLICATION NUMBER: US 09/665,791
11 PRIOR FILING DATE: 2000-10-10
12 PRIOR APPLICATION NUMBER: US 09/513,904
13 PRIOR FILING DATE: 2000-02-25
14 PRIOR APPLICATION NUMBER: US 09/507,533
15 PRIOR FILING DATE: 2000-02-18
16 PRIOR APPLICATION NUMBER: US 60/168,780
17 PRIOR FILING DATE: 1999-12-03
18 NUMBER OF SEQ ID NOS: 19
19 SOFTWARE: FastSeq for Windows Version 4.0
20 SEQ ID NO 11
21 LENGTH: 1147
22 TYPE: RPT
23 ORGANISM: Homo sapiens
24 OS-09-798-412-11

```

[illegible]

```

RESULT 7
US-10-325-917-11
; Sequence 11, Application US/10325917
; Publication No. US20050113787A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; TITLE OR INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-327001
; CURRENT APPLICATION NUMBER: US/10/325,917
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US/09/798,412
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 09/728,260
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/665,791
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 09/513,904
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: US 09/507,533
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: US 60/168,780

```

```

: PRIOR FILING DATE: 1999-12-03
: NUMBER OF SEQ. ID NOS: 19
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 11
: LENGTH: 1147
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-325-917-11

```

Query Match	37.6%;	Score 270.5;	DB 15;	Length 1147;
Best Local Similarity	48.6%;	Pred. No. 1.3e-22;		
Matches	54;	Conservative	23;	Mismatches 33;
			Indels	1;
			Gaps	1;

Qy 15 DEEHTMMESHRRIVRCICPSHPTFYLAQAKLCOLDEEEVHSHRLTMSARACHLL 74

Db 11 EEDALFENVECNRRMLSRVINPAKLTPLYLQCKVITDEQDEDEVYLAAMPALPSKINRAGRLL 70

Qy 75 DLKTRKNGAIAFLESLKFNHPVYLYLVNGLPDVDFSNFGSGSSPFDGL 125

Db 71 DILHTKQRGVVFLESLEFYLPDELVLKLVGKEPTRLRSTIIVEEG-HREG 120

RESULT 8  
US-10-032-159A-8  
; Sequence 8, Application US/10032159A  
; Publication No. US20020164703A1

```

1  APPLICANT: Pawlowski, Krzysztof
2  APPLICANT: Reed, John C.
3  APPLICANT: Godzik, Adam
4  TITLE OF INVENTION: CARD-DOMAIN CONTAINING POLYPEPTIDES,
5  TITLE OF INVENTION: ENCODING NUCLEIC ACIDS, AND METHODS OF USE
6  FILE REFERENCE: P-LJ 5100
7  CURRENT APPLICATION NUMBER: US/10/032,159A
8  CURRENT FILING DATE: 2001-12-19
9  PRIOR APPLICATION NUMBER: US 60/257,457
10 PRIOR FILING DATE: 2000-12-21
11 NUMBER OF SEQ ID NOS: 37
12 SOFTWARE: FastSeq for Windows Version 4.0
13 SEQ ID NO 8
14 LENGTH: 1247
15 TYPE: PRt
16 ORGANISM: Homo sapiens
17 US-10-032-159A-8

```

```

Query March 37.6%; Score 270.5; DB 14; Length 1247;
Best Local Similarity 48.6%; Pred. No. 1.5e-22;
Matches 54; Conservative 23; Mismatches 33; Indels 1; Gaps 1;

QY 15 DEETLMMESHSHRIIVRCICPSRLTFYLRQAKVLCQDDEEVLHSPRLTNSAMRAGHL 74
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 11 EEDLMMENVCNRMHLSRYINPAKLTYYLQCKYIDEDQDEVLNAAEMLPSKINRAGRL 70
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

QY 75 DLKTRKGNGAIAFLSLKFNHPDVTYLTGLOPDVDFSNFSGSSDPDGI 125
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 71 DILHTKGQRGVVFLESLERFYELLYLTGKERTFRFSITVVEG--HEGL 120

RESULT 9
US-10-032-159A-10
; Sequence 10, Application US/10032159A
; Publication No. US20020164703A1
; GENERAL INFORMATION:
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; TITLE OF INVENTION: CARD-DOMAIN CONTAINING POLYPEPTIDES,
; FILE REFERENCE: P-1J 5100
; CURRENT APPLICATION NUMBER: US/10/032.159A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 60/257,457
; PRIOR FILING DATE: 2000-12-21

```

NUMBER OF SEQ ID NOS: 37  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10  
LENGTH: 92  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-032-159a-10

Query Match 35.7%; Score 257; DB 14; Length 92;  
Best Local Similarity 53.3%; Pred. No. 1.6e-22;  
Matches 48; Conservative 20; Mismatches 22; Indels 0; Gaps 0;

QY 16 DEETLWEMSHRHRIVRCICPSRLTPYLRQAKVLCQDDEEVLSHSPRLTNSAMRAGHLL 75  
DB 1 EDALWENECRHRHLSTRINPAKLTPLRQCKVIDEDEDVLANAPLPSKINRAGRLLD 60  
QY 76 DLKTRGNKAIAFLSKLFHNPDVYTLVTG 105  
DB 61 ILHTKGRGVVVFLESLEFYYPPELYKLVTG 90

RESULT 10  
US-09-798-412-8  
Sequence 8, Application US/09798412  
Publication No. US20030109428A1  
GENERAL INFORMATION:  
APPLICANT: Berlin, John  
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED  
FILE REFERENCE: 07334-327001  
CURRENT FILING DATE: US/09/798,412  
PRIOR FILING DATE: 2001-03-02  
PRIOR APPLICATION NUMBER: US 09/728,260  
PRIOR FILING DATE: 2000-12-01  
PRIOR APPLICATION NUMBER: US 09/685,791  
PRIOR FILING DATE: 2000-10-10  
PRIOR APPLICATION NUMBER: US 09/513,904  
PRIOR FILING DATE: 2000-02-25  
PRIOR APPLICATION NUMBER: US 09/507,533  
PRIOR FILING DATE: 2000-02-18  
PRIOR APPLICATION NUMBER: US 60/168,780  
PRIOR FILING DATE: 1999-12-03  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8  
LENGTH: 1032  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-798-412-8

Query Match 34.2%; Score 246; DB 11; Length 1032;  
Best Local Similarity 48.9%; Pred. No. 9.2e-20;  
Matches 46; Conservative 21; Mismatches 27; Indels 0; Gaps 0;

QY 15 DEETLWEMSHRHRIVRCICPSRLTPYLRQAKVLCQDDEEVLSHSPRLTNSAMRAGHLL 74  
DB 23 EDALWENECRHRHLSTRINPAKLTPLRQCKVIDEDEDVLANAPLPSKINRAGRLLD 82  
QY 75 DLKTRGNKAIAFLSKLFHNPDVYTLVTG 108  
DB 83 DLKTRGNKAIAFLSKLFHNPDVYTLVTG 116

RESULT 11  
US-10-325-917-8  
Sequence 8, Application US/10325917  
Publication No. US20030113787A1  
GENERAL INFORMATION:  
APPLICANT: Berlin, John  
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED  
FILE REFERENCE: 07334-327001  
CURRENT APPLICATION NUMBER: US/10/325,917

CURRENT FILING DATE: 2002-12-20  
PRIOR APPLICATION NUMBER: US/09/798,412  
PRIOR FILING DATE: 2001-03-02  
PRIOR APPLICATION NUMBER: US 09/728,260  
PRIOR FILING DATE: 2000-12-01  
PRIOR APPLICATION NUMBER: US 09/685,791  
PRIOR FILING DATE: 2000-10-10  
PRIOR APPLICATION NUMBER: US 09/513,904  
PRIOR FILING DATE: 2000-02-25  
PRIOR APPLICATION NUMBER: US 09/507,533  
PRIOR FILING DATE: 2000-02-18  
PRIOR APPLICATION NUMBER: US 60/168,780  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8  
LENGTH: 1032  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-325-917-8

Query Match 34.2%; Score 246; DB 15; Length 1032;  
Best Local Similarity 48.9%; Pred. No. 9.2e-20;  
Matches 46; Conservative 21; Mismatches 27; Indels 0; Gaps 0;

QY 15 DEETLWEMSHRHRIVRCICPSRLTPYLRQAKVLCQDDEEVLSHSPRLTNSAMRAGHLL 74  
DB 23 EDALWENECRHRHLSTRINPAKLTPLRQCKVIDEDEDVLANAPLPSKINRAGRLLD 82  
QY 75 DLKTRGNKAIAFLSKLFHNPDVYTLVTG 108  
DB 83 DLKTRGNKAIAFLSKLFHNPDVYTLVTG 116

RESULT 12  
US-09-798-412-2  
Sequence 2, Application US/09798412  
Publication No. US20030109428A1  
GENERAL INFORMATION:  
APPLICANT: Berlin, John  
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED  
FILE REFERENCE: 07334-327001  
CURRENT FILING DATE: US/09/798,412  
PRIOR FILING DATE: 2001-03-02  
PRIOR APPLICATION NUMBER: US 09/728,260  
PRIOR FILING DATE: 2000-12-01  
PRIOR APPLICATION NUMBER: US 09/685,791  
PRIOR FILING DATE: 2000-10-10  
PRIOR APPLICATION NUMBER: US 09/513,904  
PRIOR FILING DATE: 2000-02-25  
PRIOR APPLICATION NUMBER: US 09/507,533  
PRIOR FILING DATE: 2000-02-18  
PRIOR APPLICATION NUMBER: US 60/168,780  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 536  
TYPE: PRT  
ORGANISM: Rattus norvegicus  
US-09-798-412-2

Query Match 31.0%; Score 223.5; DB 11; Length 536;  
Best Local Similarity 45.4%; Pred. No. 1.7e-17;  
Matches 49; Conservative 18; Mismatches 38; Indels 3; Gaps 1;

QY 15 DEETLWEMSHRHRIVRCICPSRLTPYLRQAKVLCQDDEEVLSHSPRLTNSAMRAGHLL 74  
DB 6 NDECSALSFRRVRLISVIDPSRITPYLRQCKVIDEDEDVLANAPLPSKINRAGRLLD 65  
QY 75 DLKTRGNKAIAFLSKLFHNPDVYTLVTG 119

Db 66 DILQRTGHKGVAFLFSLLELYYPOLYRKVTGKEPARVFSMTIDAGS 113

```
RESULT 13
US-10-325-917-2
; Sequence 2, Application US/100325917
; Publication No. US20030113787A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; FILE REFERENCE: 07334-327001
; CURRENT APPLICATION NUMBER: US/10/325,917
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US/09/798,412
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 09/728,260
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/685,791
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 09/513,904
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: US 09/507,533
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: US 60/168,780
; PRIOR FILING DATE: 1999-12-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-325-917-2
```

```
Query Match 31.0%; Score 223.5; DB 15; Length 536;
Best Local Similarity 45.4%; Pred. No. 1.7e-17;
Matches 49; Conservative 18; Mismatches 38; Indels 3; Gaps 1;

QY 15 DEETLWEMESHRRHRIYRCICPSRLTPYLRQAKVLCQDDEEVLHSPRLTNSAMRAGHL 74
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 6 NDECSWVLEGFRTVLTLSVIDPSRITPYLRQCKVLPDDEQVLSDPNLTIRKRVGVLL 65

QY 75 DLKTRGNKGAIAFLFSLKFNPDVYTLVTGLQPDVDFS--NFSGES 119
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 66 DILQRTGHKGVAFLFSLLELYYPOLYRKVTGKEPARVFSMTIDAGS 113
```

```
RESULT 14
US-10-032-159A-2
; Sequence 2, Application US/10032159A
; Publication No. US20020164703A1
; GENERAL INFORMATION:
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; TITLE OF INVENTION: CARD-DOMAIN CONTAINING POLYPEPTIDES,
; FILE REFERENCE: P-LJ 5100
; CURRENT APPLICATION NUMBER: US/10/032,159A
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 60/257,457
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-032-159A-2
```

```
Query Match 29.9%; Score 215.5; DB 14; Length 366;
Best Local Similarity 44.4%; Pred. No. 8.8e-17;
Matches 48; Conservative 17; Mismatches 40; Indels 3; Gaps 1;
```

```
QY 15 DEETLWEMESHRRHRIYRCICPSRLTPYLRQAKVLCQDDEEVLHSPRLTNSAMRAGHL 74
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 6 NDECSWVLEGFRTVLTLSVIDPSRITPYLRQCKVLPDDEQVLSDPNLTIRKRVGVLL 65

QY 75 DLKTRGNKGAIAFLFSLKFNPDVYTLVTGLQPDVDFS--NFSGES 119
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 66 DILQRTGHKGVAFLFSLLELYYPOLYRKVTGKEPARVFSMTIDAGS 113
```

```
RESULT 15
US-10-032-159A-20
; Sequence 20, Application US/10032159A
; Publication No. US20020164703A1
; GENERAL INFORMATION:
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; TITLE OF INVENTION: CARD-DOMAIN CONTAINING POLYPEPTIDES,
; FILE REFERENCE: P-LJ 5100
; CURRENT APPLICATION NUMBER: US/10/032,159A
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 60/257,457
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-032-159A-20
```

```
Query Match 29.9%; Score 215.5; DB 14; Length 366;
Best Local Similarity 44.4%; Pred. No. 8.8e-17;
Matches 48; Conservative 17; Mismatches 40; Indels 3; Gaps 1;

QY 15 DEETLWEMESHRRHRIYRCICPSRLTPYLRQAKVLCQDDEEVLHSPRLTNSAMRAGHL 74
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 6 NDECSWVLEGFRTVLTLSVIDPSRITPYLRQCKVLPDDEQVLSDPNLTIRKRVGVLL 65

QY 75 DLKTRGNKGAIAFLFSLKFNPDVYTLVTGLQPDVDFS--NFSGES 119
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 66 DILQRTGHKGVAFLFSLLELYYPOLYRKVTGKEPARVFSMTIDAGS 113
```

Search completed: February 18, 2004, 04:56:50  
Job time : 78 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: February 17, 2004, 21:36:16 ; Search time 1366.86 Seconds

(without alignments)  
8260.598 Million cell updates/sec

Title: US-10-032-159A-17

Perfect score: 276  
Sequence: 1 gagagagacactgtgagagat.....ccctgtcacccggcgcag 276Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hng:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vt:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vt:\*  
30: em\_hng\_hum:\*  
31: em\_hng\_inv:\*  
32: em\_hng\_other:\*  
33: em\_hng\_mus:\*  
34: em\_hng\_pln:\*  
35: em\_hng\_rnd:\*  
36: em\_hng\_mam:\*  
37: em\_hng\_vrt:\*  
38: em\_sy:\*  
39: em\_hngo\_hum:\*  
40: em\_hngo\_mus:\*  
41: em\_hngo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	276	100.0	2621	9	BC018142	BC018142 Homo sapi
2	276	100.0	3662	9	AY032927	AY032927 Homo sapi
3	276	100.0	3766	9	AK091123	AK091123 Homo sapi
4	276	100.0	3951	9	AF322642	AF322642 Homo sapi
5	213.6	77.4	3772	10	AF363457	AF363457 Homo sapi
6	213.6	77.4	3995	10	BC029102	BC029102 Mus muscu
7	167	60.5	55173	2	AC132816	AC132816 Homo sapi
8	167	60.5	144000	9	AC123764	AC123764 Homo sapi
9	167	60.5	182016	9	AC087741	AC087741 Homo sapi
10	167	60.5	187865	2	AC015559	AC015559 Homo sapi
11	167	60.5	207822	2	AC109321	AC109321 Homo sapi
12	143.6	52.0	74301	2	AC137736	AC137736 Homo sapi
13	131.6	47.7	65608	2	AC132197	AC132197 Homo sapi
14	122.4	44.3	244653	10	AL645811	AL645811 Mouse DNA
15	113.4	41.1	3096	6	AX154570	AX154570 Sequence
16	113.4	41.1	3908	9	AY032928	AY032928 Homo sapi
17	113.4	41.1	3948	6	AX300499	AX300499 Sequence
18	113.4	41.1	3949	6	AX154568	AX154568 Sequence
19	113.4	41.1	3949	9	AY028896	AY028896 Homo sapi
20	111	40.2	55173	2	AC132816	AC132816 Homo sapi
21	110.6	40.1	4085	10	AY135367	AY135367 Mus muscu
22	105.4	38.2	4438	10	AF363456	AF363456 Mus muscu
23	100.4	36.4	3441	6	AX154573	AX154573 Sequence
24	100.4	36.4	3955	9	AF352576	AF352576 Homo sapi
25	100.4	36.4	4276	6	AX154571	AX154571 Sequence
26	100.4	36.4	4276	6	AF322641	AF322641 Homo sapi
27	100.4	36.4	4293	9	AK074049	AK074049 Homo sapi
28	88	31.9	765	6	BD150612	BD150612 Primer fo
29	88	31.9	1608	6	AX154567	AX154567 Sequence
30	88	31.9	2098	6	AX154565	AX154565 Sequence
31	88	31.9	2132	9	AF311287	AF311287 Homo sapi
32	88	31.9	2176	6	BD160313	BD160313 Primer fo
33	88	31.9	2176	6	AK024001	AK024001 Homo sapi
34	86.4	31.3	1835	6	BC008877	BC008877 Homo sapi
35	79.6	28.8	1608	6	AX154564	AX154564 Sequence
36	79.6	28.8	1879	6	AX154562	AX154562 Sequence
37	79.6	28.8	1879	10	AF311288	AF311288 Rattus no
38	72.6	26.3	239392	2	AC115417	AC115417 Rattus no
39	72.6	26.3	245032	2	AC109749	AC109749 Rattus no
40	72.4	26.2	96256	9	HS117715	AL022315 Human DNA
41	69.4	25.1	187446	2	AC124374	AC124374 Mus muscu
42	68.8	25.1	225027	2	AC131725	AC131725 Mus muscu
43	68.8	24.9	3164	10	BC004692	BC004692 Mus muscu
44	67.8	24.6	161985	10	AL592169	AL592169 Mouse DNA
45	67.8	24.6	217819	2	AC026386	AC026386 Mus muscu

## ALIGNMENTS

RESULT 1  
BC018142  
LOCUS  
DEFINITION  
Homo sapiens, Similar to caespase recruitment domain protein 14,  
clone MGC:9539 IMAGE:3847282, mRNA, complete cdt.  
ACCESSION  
BC018142  
VERSION  
BC018142.1 GI:17390314  
SOURCE  
MGC.  
ORGANISM  
Homo sapiens (human)  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 2621)  
AUTHORS  
Strausberg, R.  
TITLE  
Direct Submission

## JOURNAL

Submitted (03-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

## REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)

Tissue Procurement: ATCC  
DNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www-shgc.stanford.edu>  
Contact: (Dickson, Mark) [mcdpax11.stanford.edu](mailto:mcdpax11.stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
Series: IRAX Plates: 20 Row: 1 Column: 3  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 16507953.

## FEATURES

Location/Qualifiers  
1..2621  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:9539 IMAGE:3847282"  
/tissue\_type="Colon, adenocarcinoma"  
/clone\_id="NIH MGC\_65"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
178..2400  
/codon\_start=1  
/product="Similar to caspase recruitment domain protein 14"

## CDS

/protein\_id="AAH18142.1"  
/db\_xref="GI:17390315"  
/translation="MSGLCRDLSALFALDEETLMEWESHRRIVRCICPSRLTPYLR  
QAKVLCOLDEBEVSLPRLTNSAMRAGHLIDLKTRGKNGAIAFLESIFKFNPDVTL  
VTGLQPDVDFNSFGLMETSKLTCLAGALISQELNKGKQKVEVLRFNCOOLJBL  
GLAETRAEGHLEADSRMRKREVAHFHEVLRKDBMLSLIHSYNNALQKELASR  
CRSLQELYLILKQELQKRNAYSSCELEQOSLRITASDQSGDEBELNKLKEBEXKRS  
LTFSLAEKDLILOSLEDEARSRQELVERIHSLEKRAVAERQREYWEKEQTLQFO  
KSKMAQOLYREKYNALQAVCELOKEDQASARQELISQSLVEKSLRQVLEL  
TDVCEKLTROLOQAPPGVILKQEARTRCPREKQRLVYMAHICPRDSDCLVSS  
TESGLSLDLSATSRRELVDFFRSSPAPSOGLYKRAVDGFEPPMFSQCEITEG  
DQGLPRAKAGDPHLDYELDLTADLPOLBESSLOPVSRGLDVSEGVLMRRPARIL  
SQVTMLAFQGDALBEQISVTGKNTGTFIRVTPGSADQMLRPGTQVWVDYASE  
PLFPAVLEDTTLEAVGLRRVDFCCLSVKVTIDGKRLQDLKAVATSGDSFYIR  
VNLAIEGRAGELOVHCNEVILVTDTEFCGCGCMHARVNSYTMKDTAAHGTTPNYSR

BASE COUNT 563 a 782 c 857 g 419 t  
ORIGIN

Query Match 100.0%; Score 276; DB 9; Length 2621;  
Best Local Similarity 100.0%; Pred. No. 2.9e-44;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GAGAGACACTGTGGAGATGATGAGAGACCGGACACGATCGTACCTGATTCG 60  
223 GAGAGACACTGTGGAGATGATGAGAGACCGGACACGATCGTACCTGATTCG 282  
61 CCCAGCGGCTCACCCCTACCTGCGGACGAGGCGGAGGCGGACGAGGAG 120  
283 CCCAGCGGCTCACCCCTACCTGCGGACGAGGCGGAGGCGGACGAGGAG 342  
121 GAGGTGCTGACAGACCCCGGCTCACCAACAGCGGCGGAGGCGGACGAGGAG 180  
343 GAGGTGCTGACAGACCCCGGCTCACCAACAGCGGCGGAGGCGGACGAGGAG 402  
181 TTGCTGAAGACTCGAGGAGAGACGAGGAGGCGGCTTCTGAGAGAGCTTGAAGTTTCAC 240

## Db

403 TTGCTGAAGACTCGAGGAGAGACGAGGAGGCGGCTTCTGAGAGAGCTTGAAGTTTCAC 462  
241 AACCTGACGCTTACACCCCTGTCACCGGAGTCGAG 276  
463 AACCTGACGCTTACACCCCTGTCACCGGAGTCGAG 498

## Db

RESULT 2  
AY032927  
LOCUS  
DEFINITION  
Homo sapiens CARD-containing MAGUK 2 protein (CARMA2) mRNA,  
complete cds.  
AY032927  
ACCESSION  
AY032927.1 GI:14192722  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 3682)  
Gaido, O., Martignon, F., Mischeau, O., Bonnet, D., Thome, M. and  
Tschopp, J.  
Carnal, a CARD-containing binding partner of Bcl10, induces Bcl10  
phosphorylation and NF-kappa activation  
FEBS Lett. 496 (2-3), 121-127 (2001)  
2 (bases 1 to 3682)  
11356195  
PUBMED  
JOURNAL  
MEDLINE  
REFERENCE  
AUTHORS  
MARTIGNON, F., THOME, M. and TSCHOPP, J.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (23-APR-2001) Institute of Biochemistry, University of  
Lausanne, Ch. des Boveresses 155, Epalinges, VD 1066, Switzerland  
Location/Qualifiers  
1..3682  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
1..3682  
/gene="CARMA2"  
4..3018  
/gene="CARMA2"  
/note="contains caspase recruitment domain CARD, coiled  
coil, PDZ, SH3, guanylate kinase domain GUK, interacts  
with the CARD-containing molecule Bcl10 (CARMEN)"  
/codon\_start=1  
/product="CARD-containing MAGUK 2 protein"  
/protein\_id="AAK54453.1"  
/db\_xref="GI:14192723"  
/translation="MSGLCRDLSALFALDEETLMEWESHRRIVRCICPSRLTPYLR  
QAKVLCOLDEBEVSLPRLTNSAMRAGHLIDLKTRGKNGAIAFLESIFKFNPDVTL  
VTGLQPDVDFNSFGLMETSKLTCLAGALISQELNKGKQKVEVLRFNCOOLJBL  
GLAETRAEGHLEADSRMRKREVAHFHEVLRKDBMLSLIHSYNNALQKELASR  
CRSLQELYLILKQELQKRNAYSSCELEQOSLRITASDQSGDEBELNKLKEBEXKRS  
LTFSLAEKDLILOSLEDEARSRQELVERIHSLEKRAVAERQREYWEKEQTLQFO  
KSKMAQOLYREKYNALQAVCELOKEDQASARQELISQSLVEKSLRQVLEL  
TDVCEKLTROLOQAPPGVILKQEARTRCPREKQRLVYMAHICPRDSDCLVSS  
TESGLSLDLSATSRRELVDFFRSSPAPSOGLYKRAVDGFEPPMFSQCEITEG  
DQGLPRAKAGDPHLDYELDLTADLPOLBESSLOPVSRGLDVSEGVLMRRPARIL  
SQVTMLAFQGDALBEQISVTGKNTGTFIRVTPGSADQMLRPGTQVWVDYASE  
PLFPAVLEDTTLEAVGLRRVDFCCLSVKVTIDGKRLQDLKAVATSGDSFYIR  
VNLAIEGRAGELOVHCNEVILVTDTEFCGCGCMHARVNSYTMKDTAAHGTTPNYSR

## FEATURES

source  
gene  
CDS  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
1..3682  
/gene="CARMA2"  
4..3018  
/gene="CARMA2"  
/note="contains caspase recruitment domain CARD, coiled  
coil, PDZ, SH3, guanylate kinase domain GUK, interacts  
with the CARD-containing molecule Bcl10 (CARMEN)"  
/codon\_start=1  
/product="CARD-containing MAGUK 2 protein"  
/protein\_id="AAK54453.1"  
/db\_xref="GI:14192723"  
/translation="MSGLCRDLSALFALDEETLMEWESHRRIVRCICPSRLTPYLR  
QAKVLCOLDEBEVSLPRLTNSAMRAGHLIDLKTRGKNGAIAFLESIFKFNPDVTL  
VTGLQPDVDFNSFGLMETSKLTCLAGALISQELNKGKQKVEVLRFNCOOLJBL  
GLAETRAEGHLEADSRMRKREVAHFHEVLRKDBMLSLIHSYNNALQKELASR  
CRSLQELYLILKQELQKRNAYSSCELEQOSLRITASDQSGDEBELNKLKEBEXKRS  
LTFSLAEKDLILOSLEDEARSRQELVERIHSLEKRAVAERQREYWEKEQTLQFO  
KSKMAQOLYREKYNALQAVCELOKEDQASARQELISQSLVEKSLRQVLEL  
TDVCEKLTROLOQAPPGVILKQEARTRCPREKQRLVYMAHICPRDSDCLVSS  
TESGLSLDLSATSRRELVDFFRSSPAPSOGLYKRAVDGFEPPMFSQCEITEG  
DQGLPRAKAGDPHLDYELDLTADLPOLBESSLOPVSRGLDVSEGVLMRRPARIL  
SQVTMLAFQGDALBEQISVTGKNTGTFIRVTPGSADQMLRPGTQVWVDYASE  
PLFPAVLEDTTLEAVGLRRVDFCCLSVKVTIDGKRLQDLKAVATSGDSFYIR  
VNLAIEGRAGELOVHCNEVILVTDTEFCGCGCMHARVNSYTMKDTAAHGTTPNYSR  
A000LIALIOMTQCTVTRKPSGGPQKLVIVSMKAKASPLSLFDRQDLSRM  
EGSSCFWASCLTLPYTLVWPRPARPRVILYRAVAGLISKLCLLOGFKKLA  
EYLSQEEYVANSQGRDIIORGEVGGRCWYTRHVESIMKNTALLDVOLDVCTTH  
RMDIPRTIYVSNVERAKKIKKGLQRTGBBQILFARQRBQDLDPAFCIYSLAP  
DQMSLDLILSCVRQALADBEQKRVWTEQSPR"

## BASE COUNT

751 a 1100 c 1194 g 637 t  
ORIGIN

Query Match 100.0%; Score 276; DB 9; Length 3682;  
Best Local Similarity 100.0%; Pred. No. 2.8e-44;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;